AN INVESTIGATION OF A READING MODEL FOR ENGLISH AS A SECOND LANGUAGE

WILLY D. ENGINEER

Thesis submitted in fulfillment of the requirements for the degree of Ph.D. in Linguistics in the University of Edinburgh

1977

ABSTRACT OF THESIS

The thesis represents an experimental investigation of the reading process *via* a series of task-induced operations involving the use of long texts in a second language. The theoretical construct for the operations is provided by a reading model which has been developed to integrate not only known facts of perception already established through experimental findings, but also psychological explanations of processes we have hypothesised must be involved if graphic material in a second language is to be interpreted meaningfully.

Based on earlier published models, the five-stage model describes the components, processes and limitations of each stage of the reading activity. And because it fulfills a primarily functional role, the model includes the output of responses that permit the efficiency of the cognitive operations to be evaluated through observable behaviour.

The experiments are aimed at eliciting information regarding both the process and the product of comprehension. Techniques suited for these specific ends are discussed along with other task variables that might affect performance. Of special interest are the alternative methods suggested for analysis of results. Data from two informal and four formal experiments provide support for the validity of the reading model, while the use of long texts has produced fairly convincing evidence that length is a factor contributing to discrimination between levels of reading ability.

Total responses for the process experiment were analysed to develop what we have called the Reading Performance Profile. This represents patterns of performance based on correct, incorrect and omission scores that can be identified with levels of reading ability. We discuss the use of this profile as a diagnostic tool along with implications of our findings which have special relevance for the non-native speaker population used in our investigation.

PREFACE

There are five sections to this thesis.

Section One comprises the first two chapters which together provide the necessary background information for the investigation. It includes details regarding the aim and purpose of the study along with information concerning the non-native speaker population we were interested in. Definitions of some of the terms that help delimit the parameters of the investigation are also included.

Section Two is primarily a review of the relevant literature. In Chapter Three we present a brief survey of some of the standardized tests of proficiency in English as a second language in use today at universities in the United States and Britain, and report various experiments in reading from comparatively recent literature in areas of interest to us. Chapter Four presents a critical survey of some of the better known models of reading, and discusses four of these in greater detail along with the implications for our model.

Section Three presents our five-stage model of reading. In Chapter Six, the specific hypotheses postulated on the basis of the working model are formulated, and the other task variables involved in the testing situation are discussed. These include the materials to be used in the experiments, the techniques selected for the testing, and the scoring procedures to be adopted. This section ends with a design of the four formally planned experiments.

Section Four presents the results of the experiments.

Section Five comprises the final three chapters of the thesis. Chapter Nine discusses the relevance of the results to the hypotheses investigated. Also developed is the Reading Performance Profile, a diagnostic instrument aimed at providing information regarding the subject's reading efficiency. Applications of the findings are discussed as are suggestions for further research. Chapter Eleven summarises the findings and presents the concluding remarks of the thesis.

DECLARATION

This thesis is my own work and composition.

Willy D. Engineer
8 November, 1977.

ACKNOWLEDGEMENTS

My thanks go initially to The British Council without whose generous offer to finance my studies this research would not have been possible. In this connection I would like especially to mention my colleagues and friends in the Bombay office of the Council, and members of the Staff Training Department in London for their continued kindness and encouragement during these three years.

In addition I owe a lot to my family for taking over responsibility of my personal affairs leaving me free to undertake this study.

The pragmatic nature of this thesis inevitably owes a great deal to all my many friends who co-operated in conducting the experiments, and although these are too numerous to list, I would like to single out a few : R. Vania and J. Donoghoe at Krefeld and D. Heath at Nuremberg in West Germany, P. Wingard of Manchester University at Alexandria in Egypt, C. Wallace and R. Hornsey of Ealing College of Technology, E. Glendinning of the English Language Foundation at Haddington, J. Trench and J. Mahood of the Bell School of Languages at Norwich. My thanks go to all the students at these various centres and here at Edinburgh University who gave so generously of their time in sitting the tests.

More than most, my gratitude is due to my supervisor Dr. Alan Davies, not just for his patience but more importantly, for persistent questioning regarding the practical issues involved in the investigation thereby ensuring that my feet stayed on the ground.

Lastly my thanks to Mrs. M. O'Donnell, and particularly Mrs. S. Motherwell for typing this thesis in record time.

CONTENTS : VOLUME I

LIST OF FIGURES

LIST OF TABLES

SECTION ONE

BACKGROUND INFORMATION

CHAPTER ONE

INTRODUCTION : AIM AND PURPOSE

In this thesis I seek to present a theoretical model of the reading process when reading in a second language in order to provide a theoretical basis for specified tasks that required the non-native speaker of English to produce evidence of his understanding of on-going discourse. These tasks were aimed at investigating and distinguishing between the process of reading and the end-product of the activity, since we felt that each had a contribution to make to our understanding of the complex of skills that are at work when reading is successfully undertaken. Because our interest was with non-native speakers who for academic reasons needed to be efficient in reading, the materials used to test their ability comprised longer texts than are generally found in foreign language course books or in tests of English as a second language.

It was accepted that non-native speakers coming from various linguistic backgrounds would possess differing levels of ability in the foreign language which for our study was English. They were expected now to attend courses at university level in English, and needed to possess the ability not only to recognise language forms and the cognitive relations they were meant to represent, but also to appreciate their value in the context of use.

It was hoped that this investigation would throw more light both on the techniques and the tasks used to elicit evidence of comprehension and reading ability.

The study therefore investigated performance in specific tasks based on long texts, that reflected the theoretical model proposed in order to ascertain the degree to which the tasks succeeded in discriminating between different levels of reading proficiency.

The writer is only too well aware of the limitations in our knowledge of the reading activity. What is feasible is the setting up of tasks which are adjudged reliable and valid measures of the skill being evaluated. By requiring the production of responses that can be matched against a criterial set that have been predetermined as indicative of successful completion of the task, one can infer the degree of reading skill that is possessed by the subject.

It was hoped that in the analysis of results, a recognizable pattern of responses might emerge that could be used to identify levels of proficiency in reading. With the development of a 'reading performance profile' based on the responses generated by the tasks, the strengths and weaknesses could be highlighted and subsequently utilised for diagnostic purposes.

An important part of the study was hence devoted to the development of reliable and valid measures of reading ability. These then permitted us to determine the level of proficiency possessed by the non-native speaker when reading in a second language.

CHAPTER TWO

BACKGROUND INFORMATION

2. GENERAL REMARKS

Over the last fifteen years there has been a vast proliferation in numbers of non-native speakers of English enrolling for higher education courses at institutions where English is the language of instruction.

These institutions are in some cases in the student's own country, while in others they are situated in an English speaking environment. Although we were primarily concerned with students in the latter context, because of the relevance of this investigation to the overseas situation, three sample groups representing the former category were also included in our study.

This investigation then focussed on the reading ability of students admitted to academic institutions in Britain. Foreign students are admitted to academic institutions on the basis of entrance qualifications dictated by the various courses of study for which they are enrolled. In this, they have competed, presumably with native speakers of English, with one proviso. Many universities today, both in Britain and the United States, require non-native speakers - (NNS) - to produce some proof of proficiency in English to satisfy the authorities that they possess sufficient knowledge of the language to pursue their studies effectively.

What constitutes a 'sufficient knowledge' varies from one institution to another, but is generally determined by success in a test that is accepted as providing a valid measure of language proficiency. These tests usually sat in the candidate's own country vary from subjectively assessed interviews to standardised batteries. Results regarding the candidate's linguistic ability are generally included along with application forms and are used for offering provisional admission to the university.

But because of the wide variation in reliability of the measures used, an increasing number of university authorities require all foreign students at matriculation to sit a test in English administered by the university to which they are admitted. These tests are not identical in form and content, and Chapter Three includes a brief survey of some of these. On the basis of these tests, students are often advised to attend remedial courses. What is of interest however is that admission is seldom refused or withdrawn despite poor performance in the proficiency test.

A number of NNS are scholarship holders of one sort or another, and many arrive in Britain some months before term begins. This enables them to enroll on English language courses, many of which are specially designed to help the prospective students overcome language deficiency. However since the obvious need of most foreigners on arrival is ability to communicate intelligibly during face to face interaction, this is the area of concern

that is most often attended to. Private language schools are reasonably well-equipped with audio equipment including language laboratories, and much of the curriculum time is devoted to remediating this aural/oral weakness.

Relatively little attention is given to the development of reading and writing skills which may well turn out to be as crucial as the aural/oral ones in the successful completion of a course of study at the university. The necessity of having to live in a new linguistic community results in a fairly impressive improvement of oral communication skills but not necessarily at a level judged to be grammatically correct.

However the demands made on their linguistic abilities for academic purposes are likely to involve both the understanding of spoken and written English, and the production of evidence that they have not only understood but are capable of contributing to an advancement of knowledge in their own specialist fields. Since the NNS are not likely to receive further tuition in reading, it seems specially relevant that we make as fair an assessment of their linguistic proficiency as possible in the students' own interests. Therefore one of the major concerns in our study was how best to elicit evidence of reading ability, and how best to measure it.

2.1 THE READING TASK

Our population was expected to possess independent reading ability with long stretches of text at a fairly advanced level. We wished then to examine the skill with

which the NNS handled the long passage and the ease with which he extracted the information content in response to specific demands imposed by the task set. For this we needed to develop reliable and valid measures that permitted us to investigate the reader's reactions to meaningful verbal discourse - MVD - (Carroll, 1971:1) while the reading was still in progress as compared to what was understood from the finished product. Our interest lay both in the process and the end-product of reading.

If various 'task-induced operations' Frederiksen (1972:211) are set up, the degree to which the task succeeds in eliciting evidence of ability, suggests how far it has been a good representative of all possible tasks that could have been set. Just as the tester chooses a task that according to his construct-based hypothesis, best or most economically provides an opportunity for the reader to exhibit those cognitive processes the experimenter wishes to activate, the subject, in this task-induced operation, replies by performing to the best of his competence within the limits set by various uncontrolled factors that may affect his performance. It is axiomatic that one cannot perform better than one's competence, but one can perform considerably lower.

To the extent the subject does well in terms of the task, he is assumed to possess a level of proficiency that the experimenter infers is 'sufficient' to see him through his academic course; to the extent he does less well than is acceptable, he is inferred to be deficient in the ability measured.

Because of the complex nature of the ability being investigated, it was imperative that various factors involved in the operation be taken into account.

2.2 FACTORS AFFECTING COMPREHENSION.

It is because comprehension depends on so many factors that one can approach the study of the reading activity from different angles, each of which provide insights that help to illumine the entire interaction.

We propose then to open our discussion by an examination of what is involved in a reading task and what therefore our non-native speaker is expected to be capable of comprehending at this advanced level. This will then provide the background necessary for the survey of proficiency tests and experiments in reading that follow in Chapter Three to be seen in their proper perspective.

Since our interest lies in meaningful verbal discourse, we will follow our discussion of factors involved in comprehension with a brief comment on how terms we introduce like 'discourse', 'comprehension', 'text' and 'long passage' are to be interpreted in the thesis.

2.2.1 Communication via participation.

Jakobson (1973) talking about speech says that speakers and hearers each make moves to ensure that intended messages are in fact transferred. The speaker uses devices like repetition, paraphrase and circumlocution among others, while the hearer has the option of making interjections and comments by way of feedback. The listener in ordinary conversation can signal non-comprehension by use of either

verbal or non-verbal devices at his command. The onus for the transfer of the message rests on the speaker.

However as the situation becomes more formal, it resembles the reading situation where opportunity for feedback is minimal, so that the onus of responsibility for comprehension now falls on the hearer/reader. This resembles the normal situation in a test where questions are posed at the end of the written communication to elicit evidence of understanding.

We have then the option of examining either the 'process' stage when the text is still only partially read and understood, or the 'end-product' stage when reading has been completed, or both. Since we were interested in both process and product of the reading activity we approached the problem using techniques to evaluate both stages in the belief that each provided complementary information contributing to our knowledge of reading ability.

The reader is expected to play an active role in the reading activity if he wishes to comprehend not just what is printed, but more importantly, what the writer meant to convey. And the probability of inferring the latter correctly is a product of what Garfinkel (1967) has called 'shared expectations'.

According to Cicourel (1973) the various interpretive procedures needed for comprehension are not on display in the utterance but are implicit in the assumptions necessary for the discourse to continue. These procedures he defines as:

(i) reciprocity of perspectives : which refers to the basic assumption that discourse participants share the same experience of the immediate scene so that if the roles were reversed, each participant could reconstruct the other's intentions.

(ii) the et cetera assumption : which permits one to be vague, and where the hearer/reader defers judgement until more is known rather than refusing to participate and thus cause a breakdown in the act of communication.

(iii) retrospective-perspective sense of occurrence : which enables one to see the continuity and coherence in the discourse, and normalise the event and tolerate the vagueness often present in descriptions which serve as indexical expressions.

All these procedures are relevant for the NNS reading in English who is expected to realise what has been deleted because it is presumed to be known to the reader. The author may omit information, delay its inclusion, or refer to it obliquely, once again because of his assumption that the reference will be easily recognised or alternatively that it can be easily recovered from the context. One of the main problems of the NNS is not just what the writer is saying but what he is talking about at a particular time. Although the general theme providing coherence to the passage may not change, the particular topic of a sentence or paragraph changes as discourse develops. The reader in order to

comprehend must be capable of ascribing the incoming information to the correct referrent.

'Efforts to put some sensible construction on what another person is saying are usually aided by knowledge of the context in which he says it. The context provides a pool of shared information on which both parties to a conversation can draw. The information, both contextual and general, that a speaker believes his listener shares with him constitutes the cognitive background of his utterance. Information that he adds constitutes the focus of his utterance, and his sentence will usually be phrased in such a way as to make it clear that this new information should be the focus of the listener's attention' Miller and Johnson-Laird (1976:125) 'Understanding is antecedent to verification, not a consequence of verification' (Ibid:126).

To summarise what has been said, the reader must comprehend both what is said and what is meant, and the interrelationship between the two; he needs to actively participate contributing his knowledge of the linguistic and social conventions that surround the acts of reading and writing to enable him to comprehend not just the dictionary meaning of a series of words that make up a text, but the value each has in the specific context, that is, the discourse function of each unit under investigation.

2.2.2 The interpretation of discourse.

'The interpretation of discourse, then, is not simply a matter of recovering the presuppositions attaching to individual sentences as they appear in sequence. The

linguistic context of utterance, create presuppositions of a pragmatic kind or "implicatives" which can over-ride those which are associated with linguistic forms', Widdowson (160)[1] And in the same paper Widdowson proceeds to make the point that 'discourse is not simply a patchwork of pre-ordained sentential meanings, it is a dynamic process of creation'.

The reader then brings to reading not only his total knowledge of the foreign language system and of the subject matter, but also of the rules of discourse. If he has not abstracted the latter as they function in English, he will perforce utilise his knowledge of these rules as they behave in his first language. To the extent that his familiarity with their use in English enables him to read and communicate intelligibly, he will be able to compensate for the presence of any ambiguity in the act of communication. Where this ambiguity may well interfere with understanding will be when the foreign reader takes a literalistic view of the language used. Context of actual use constrains the seeming polysemy of discourse, and failure to take cognizance of the various features involved in the situation may well raise barriers to understanding not expected by the writer.

We took then as our definition of discourse, Widdowson's approach that it can be 'roughly defined as the use of sentences' (1973:69) or 'as the manner in which linguistic

---

1. Approaches to Discourse in Med Ziffern (forthcoming) pp.151-167.

elements function to communicate effect' which involves 'correlating the meaning of the linguistic item as an element of the language code, with the meaning it takes in the context in which it occurs', (1975:33). Discourse is then concerned with how the language is actually used and rules of use must be learned without which knowledge of grammar alone is insufficient.

The NNS who acquires skill in the mechanical manipulation of elements of the language will not be able to use them to communicate meaning unless he also applies the rules of discourse. Nor are the two necessarily acquired sequentially without due attention paid to them. The elements of a 'discourse bloc' to use Kaplan's term (1972) may or may not coincide with sentences, but each element is a part of the whole and the functional interrelationship is both a product of the syntax and the semantics, and the position the part occupies in the discourse environment. Our readers need to appreciate how the language of the discourse bloc functions in the particular instance, and how and why the writer has used it. It is in terms of its communicative functions rather than the linguistic form that it needs to be understood.

The student is expected now to handle forms of discourse that embody highly complex structures and vocabulary. He must identify similarity or equivalence in meaning despite variations in surface forms while at the same time appreciating that underlying meanings could be very different even when superficially alike structures were used. The NNS at this

level is expected not only to have had language experience but also familiarity with cultural norms and conventions that determine how linguistic forms are to be interpreted in normal usage. This degree of familiarity will enable him to recognise deviations and variations used by the writer to focus attention on information he considers important.

This common knowledge is taken for granted by the writer who directs his work at an audience with whom he shares language, world experience and ways of thinking and who does not expect the forms he uses to be always taken literally at face value.

Items of grammar learnt in the classroom must be seen as cues to meaning and the redundancy as an in-built safety device which enables the reader to interpret accurately. The NNS must expect to come across functional varieties of a language learned in an artificial situation. Recent studies with native speakers of English, Baratz (1969), have shown the disadvantages they can suffer from if they do not possess these varieties. Although allowances are almost always made for the NNS in spoken communication, this same benefit does not extend in written form at academic level. The NNS must now compete with the native speaker in the latter's own linguistic environment, and any weaknesses in the foreigner's knowledge of the language are bound to be magnified by the manifold demands imposed on his resources from all sides.

Hence the necessity of ensuring that he possesses

a level of linguistic competence sufficient to successfully undertake a course of studies in the second language. Having seen what is expected of the NNS who reads written discourse, let us define what is meant by 'comprehension'.

2.3 COMPREHENSION.

An early definition held by many teachers and implied in much of the literature on reading in the fifties and early sixties is best expressed by Elkonin (1963) and quoted by Wayne Otto (1972:38)

'Reading is a reconstruction of the sound form of a word on the basis of its graphic representation. Understanding, which is often considered as a basic content of the process of reading, arises as a result of correct re-recreation of the sound forms of words. He who, independently of the level of understanding of words, can correctly recreate their sound form is able to read'. As can be expected a definition that introduced a dichotomy between reading of words correctly and understanding what they meant was totally unacceptable to us. The view that these were two sequential processes, with meaning explained as the result of pronunciation was out of date.

Bormuth (1969a:50) described the reading ability as 'a set of generalised knowledge-acquisition skills which permit people to acquire and exhibit information gained as a consequence of reading printed language'. Feeling that this view did not provide precise information which could be used to evaluate the ability, he re-defined the term in an attempt

to introduce a practical approach to comprehension testing, and suggested that it was the 'ability to respond correctly to a wh-question which deletes one of the immediate constituents of a syntactic structure' (1970)[1]. This definition while making for ease of constructing comprehension questions is limited to the testing of information explicitly stated in the test and hence was of little value to us.

So for our purposes we combined the views presented by Bond and Tinker who were concerned with the process, and Frank Smith who was interested in the function of the reading process.

Bond and Tinker (1957:19) stated that reading involves 'the recognition of written or printed symbols which serve as stimuli for recall of meanings built up through the reader's past experience. New meanings are derived through manipulation of concepts already in his possession. The organisation of these meanings is governed by the clearly-defined purposes of the reader. In short the reading process involves both the acquisition of the meanings intended by the writer, and the reader's contributions in the form of interpretation, evaluation, and reflection of these meanings'.

In view of the various factors affecting comprehension and the expectations created by the writer and shared by the reader to which we have already referred, this definition seemed particularly <u>apropos</u>. It also seemed particularly relevant for our NNS who already possessed a vast storehouse of concepts and knowledge which if activated could be utilised

---

1 Barmuth et al (1970)

in comprehending what he now had to read in a foreign language. We favoured the attitude that reading forms part of the overall process of knowledge acquisition, with information being added to what was already known rather than viewing it as an isolated act.

With this we combined Smith's view (1971:185) that the major aim and function of the process of reading was to view 'comprehension - or the extraction of meaning from text - as the reduction of uncertainty'. Both these views will contribute to our reading model and will be discussed in greater detail in Chapter Five. These definitions had the advantage that they contained within themselves some indication of the processes that were involved and which helped us formulate the reading model that had to be capable of performing the operations enumerated.

2.4 DEFINITIONS OF OTHER TERMS.

It remains for us to specify what we mean by some other terms used in our discussion.

2.4.1 'Process'

We have talked throughout of the reading activity as a 'process'. This concept was here seen as 'any phenomenon which shows a continuous change in time or any continuous operation or treatment' Berlo (1960:23). When we label something as a process, we also mean that it 'does not have a beginning, an end, a fixed sequence of events. It is not static, at rest. It is moving. The ingredients within the process interact, each affects all the others.

Communication theory reflects a process point of view'.(Ibid).

When one investigates a process, one is forced to arrest the motion so that the dynamic activity is interrupted. We also impose an order of occurrence for convenience so that an attempt at description is necessarily incomplete working with a fixed sequence and possibly a distorted perspective. But it is hoped that the effort is well worth while since it enables integration of the various strands of the activity into a coherent whole.

2.4.2 'Text' and 'Passage'

Two terms we have used interchangeably are 'text' and 'passage'. Halliday and Hasan (1976:1) have used the word 'text' to 'refer to any passage that forms a unified whole', and Halliday (1975:10-11) has defined it more carefully in the following way. 'Let us use the word "text" to refer to any instance of language that is operational (as distinct from citational) : everything that is said, or written, in some living context of use. Let us then conceive of text as choice. The text represents a selection within various sets of options; what is said presupposes a background of what might have been said but was not. In linguistic terms, it presupposes a paradigmatic environment: the choice of this implies the possibility of this and the other. These options are options in meaning; since our interest is in meaning rather than in the wording, we can conceive of text not as "what is said" but as "what is meant", still in the environment of what might have been - "what might have been meant (but was not)"..... A text represents a pattern of selection within a meaning potential'.

We have therefore used the words 'text' and 'passage' synonymously, and all the materials used in our experiments were self-contained texts that had to be understood as discourse if they were to communicate the information they contained.

2.4.3 The 'Long' Passage

Having stated what we mean by comprehension, discourse, process, text and passage, we need only specify how we interpreted the adjective 'long' in this study and our reasons for the decision.

The average length of passages normally used for testing reading comprehension in proficiency tests is around 150 words (see Chapter Three), and the conventional testing technique is the multiple-choice question where either the 'best' or the 'correct' alternative is to be selected from an option of four or five. The total number of words with three to four short texts is about 600 words with a time limit of 20 minutes.

But a paragraph that lends itself to three or four questions of the multiple-choice type must contain not only enough information to enable the questions to be posed but, in addition to the answers, it must enable the tester to provide efficient distractors. And in order to be efficient, distractors must contain enough of a seeming truth value to attract the less proficient reader.

Texts have therefore been chosen specifically with these requirements in mind, and with standardised tests, texts

have been discarded that failed to meet these conditions. This means that proficiency batteries use texts that contain a wealth of detail and little or no information redundancy. In terms of the test, the passage is economic and efficient. The reader must utilize his intensive study skills to evaluate the options offered as answers to the questions. No part of the passage is 'wasted'.

Moreover since the writer of a standardized test is at pains to ensure that no section of his target population is given undue advantage over another, he will choose texts that no one is familiar with. Each text will contain different information and all of this information will be relatively important. The propositional content will be high.

While it is true that the questions themselves may aim to test skills as different as direct reference, factual knowledge, inference and evaluation, the total amount of verbal material to be processed is limited by the length of the passage. Because of the quantitative restriction, the reader does not need to exercise his reading skills and strategies to re-encode items of information for later use to any great extent. The subject is not given the opportunity nor indeed is there the necessity to process the material as he reads in order to discard some of the content and more to the point, select some for preservation and later use.

Di Vesta and Gray (1973) reported that in an experiment using extended texts, more ideas were recalled in a delayed retention task when note-taking was not permitted than when recall followed immediately after the reading had been completed.[1] The implications clearly were that subjects taking part in the experiment were utilising various cognitive strategies. Processes were at work organizing the input for storage in memory. The authors went on to report that contents of a series of short segments on different topics were recalled better than the contents of a long segment of equivalent total length on a single topic. But the latter kind of reading is precisely the type that non-native speakers attending academic courses are expected to do. It was obvious that length was an important factor to be taken into account in any investigative study of this type.

The student who attends a course of instruction is expected to undertake independent reading with a fair degree of efficiency if he is to cope with the demands on his time. He must be capable of processing long stretches of running text extracting what he needs to remember. What was needed for the experiments was textual material of a length closer to the expected reading than represented by the short passages used in proficiency tests, and more importantly, on one topic. The passages needed to be too long for short term memory alone so that there was not just the

---

1. See also Bartlett's (1950) experiments with repeated productions.

opportunity but indeed the necessity of 'remembering' some of the information content while perhaps forgetting details dealt with in localised portions of the text. If the reader is to comprehend the whole, some of the facts will need to be re-encoded from a short term to a long term store for comparison and evaluation when necessary. But with a longer text, the parts that may need to be compared do not necessarily come in ordered sequences, and what is stated in one part may be modified later.

As we saw in our discussion of what was involved in a reading task, the long passage as indeed the short one, begins with certain presuppositions and assumptions that are tacitly made between writer and reader and from where the textual information proceeds. With a condensed text there is little or no opportunity for the reader to make up any deficiencies in pre-knowledge that may exist at the start. We know from the theory of communication that for a proficient reader, reduced redundancy or interference does not seriously affect comprehension in their first language.

The long text on the other hand does provide the opportunity for deficiencies to be compensated for as the passage proceeds. Information which was new at the start is itself the foundation for further new content, and concepts introduced at one stage themselves contribute to the development of ideas at a later stage.

Another important consideration that prompted us to prefer the longer text was that in most language

course books, extracts chosen to illustrate teaching points are generally short. Even when students are considered proficient or advanced enough to read extracts taken from 'literary' sources, the classroom approach often involves the breaking up of the passage into segments that can be dealt with in one period at a time. Intensive attention to detail means the inculcation of study skills so that extensive reading ability is often neglected.

It was for all these reasons that we chose to use longer texts and for our experiments interpreted the word 'long' as referring to a text of between 1000 to 1200 words which we felt was for testing purposes a viable length, while still being markedly different from the 150 word text used in proficiency batteries.

2.5 SUMMARY

In this chapter we have given the background information both of the investigation and the population with whom we were concerned. We discussed what was involved in the reading task at this advanced level and therefore the extent of knowledge and skill expected of the non-native speaker to enable him to pursue his studies effectively.

We defined the terms to be used and these by inference reflected our interest in the cognitive processes involved in reading discourse for meaning.

We have not subscribed to the alternative view that reading is a combination of various sub-skills that could be isolated for both teaching and testing and where the reading activity is viewed as an 'assembly model' Guthrie (1973). We have preferred to regard reading as a complex interrelation of many factors coming together from different sources, some from the printed verbal symbols used as stimuli, some from rules of grammar, others from rules of use and discourse. To these are added strategies and skills already possessed by the reader before he approaches the reading and which through the specific task-related objectives demanded at any time, constrain the performance of the reader.

In the following Chapter we will first briefly review some of the entrance tests of proficiency in English as a second language administered by universities, and then follow this up with a survey of some of the more interesting experiments in reading that reflect approaches similar to ours.

SECTION TWO

REVIEW OF THE LITERATURE

Chapter Three : A survey of tests and experiments in language ability.

Chapter Four : Theories and models of reading : a survey.

CHAPTER THREE

A SURVEY OF TESTS AND EXPERIMENTS IN LANGUAGE ABILITY

3. GENERAL REMARKS

We shall begin our survey by taking a brief look at some of the tests of proficiency in English which are used by universities to assess the language ability, or more specifically, the reading comprehension ability of non-native speakers; and after that to extend our interest to some of the experimental studies.

There are a number of standardised tests in use which claim to be valid measures of the language proficiency of foreign students applying for admission to academic institutions. Some few are tests designed specifically to test reading ability while the majority include a sub-test on reading comprehension as part of a wider battery. The latter word is not used here in the sense of a number of tests all aimed at measuring the same skill. Rather it refers to different sub-tests that measure different skills, which together represent proficiency. Scores from the sub-units are pooled into a composite mark so that strengths and weaknesses that may exist in individual components are balanced out. What is used for evaluation is the total score.

3.1 TESTS OF READING COMPREHENSION.

We shall begin by taking a brief look at three tests designed specifically for evaluating reading ability.

3.1.1 The Reading Comprehension Test, designed by Wiseman and Wrigley (1959) consists of 11 passages ranging in length from a single sentence to a paragraph of about 100 words. There are 60 questions using different techniques : some in multiple-choice format, some open-ended requiring one-word or short answers, and one question requiring a choice of the best summary. The passages are said to be graded, becoming more difficult with increase in length. The test was designed for students, both native and non-native, of Manchester University.

3.1.2 In 1959, the English Language Services Inc., prepared a test specially for Costa Rican students wishing to enter American universities. This too consisted of passages ranging from the single sentence to a paragraph, with 50 items in multiple-choice format. Although not designed as a speed test, the time allowed was 30 minutes on the average.

3.1.3 Of interest is the 1973 test prepared by the Pathway Industrial Unit at Ealing to assess the English language ability of ancilliary staff in hospitals. Because of an influx of immigrant workers in many of the jobs, a test in reading aloud was developed. The units ranged from the single word to sentences of varying lengths. The assessment of candidates was made by judges marking subjectively on a three-point scale.

There are of course many tests of reading used in schools, but they are generally concerned with the learning stages and so of little interest to us.

3.2 UNIVERSITY ENTRANCE TESTS : AMERICAN GROUP

Of the entrance tests in use by universities, we will briefly review nine of them, four American and five British, the latter developed mainly after 1969. Their contents are reviewed in summary form.

3.2.1 The Michigan Test of English Proficiency (1969) includes:

(1) Grammar - 40 items in multiple-choice format

(2) Vocabulary - 40 items in multiple-choice format

(3) Reading Comprehension - 20 items in multiple-choice format covering 4 passages, each of about 150 words. Time taken is 75 minutes.

In addition there are two other sub-tests:

(4) Written composition needing 30 minutes

and (5) Aural comprehension rated by judges via an interview.

3.2.2 The American Language Institute, Georgetown University test

developed in the '60's, ALI/GU, includes:

(1) English usage - 100 items in multiple-choice format to be done in 60 minutes.

(2) Oral rating - a ten to fifteen minute interview rated by 2 judges on a 5 point scale. Spoken ability is marked for general speed and length of sentences used.

(3) Listening comprehension - 40 items in multiple-choice format to be done in 15 minutes.

(4) Vocabulary and Reading - 60 items in multiple-choice format distributed between

(a) vocabulary - 25 items

(b) idioms - 10 items

and (c) reading comprehension - 25 items

covering 5 passages of 100 to 200 words.

This sub-test is to be done in 60 minutes.

3.2.3 The Test of English as a Foreign Language (1961) TOEFL administered by the Educational Testing Service, Princeton, included:

(1) Listening Comprehension - in multiple-choice format

(2) Control of English Structure - in multiple-choice format

(3) Vocabulary - in multiple-choice format

(4) Reading Comprehension - in multiple-choice format

(5) Writing ability

This battery was revised in 1976 and now includes:

(1) Listening comprehension

(2) Reading comprehension and vocabulary

(3) Structure and written expression

More important however is the reporting of partial scores along with the total obtained by the candidate. This means that the test functions not only as a proficiency instrument but for diagnostic purposes as well, being used by the supervisors tutoring on the remedial courses organised by the university.

3.2.4 The University of California, Los Angeles Placement Test (1973), UCLA, ESLPE (English as a Second Language placement test) includes:

(1) Grammar - 40 items in multiple-choice format

(2) Cloze, testing for knowledge of two-part verbs

(3) Dictation - two paragraphs 'at a fairly fast clip' Oller (1973)

(4) Reading comprehension - a number of passages with questions asking for identification of the main theme

(5) Listening comprehension - a series of dialogues with multiple-choice questions

Time taken : 2 hours

All the four university tests form batteries aimed for the most part at discrete-point testing, where various items of linguistic usage are isolated and incorporated into items for measurement. An attempt at the more 'global' or 'integrative' type of testing is evident with the introduction of cloze, dictation, composition and interviews. Most of the reading comprehension questions reflect taxonomies of the kind first made popular by Bloom (1956) and later by Barrett (1968)[1] having some questions aimed at factual information, some at inferential ability, etc. This aspect will be discussed in greater detail later.

3.3 UNIVERSITY ENTRANCE TESTS : THE BRITISH GROUP.

Let us however now take a brief look at the British tests which do much the same job as the American group.

---

[1] Quoted by Clymer (1968)

3.3.1 The English Proficiency Test (1969) developed in two parallel forms for use at Manchester University includes:

(1) Vocabulary - 50 items in multiple-choice format
Time taken : 20 to 30 minutes

(2) Grammar and usage - 50 items in multiple-choice format
Time taken : 10 minutes only

3.3.2 The English Test (1972) prepared by Leeds University includes:

(1) Grammar - 40 items in multiple-choice format

(2) Transformation - 10 items requiring production of structural change by the candidate

(3) Vocabulary - 40 items in sentence-completion type format

(4) Reading comprehension - 10 items in multiple-choice format covering 5 short passages

(5) Auditory comprehension - 25 items in multiple-choice format based on a 15 minute video-recording during which candidates are allowed to take notes.

Time for the test as a whole : 90 minutes

3.3.3 Birmingham University Diagnostic Test has been in use since 1973 and comprises:

130 items of grammar in multiple-choice format.

Time allowed is 45 minutes with an additional 15 minutes granted if requested by the candidate.

3.3.4 The English Proficiency Test Battery (EPTB) Version C. (1973), and now in a fourth form, includes two parts:

Part I

(1) Discrimination of individual sounds - 58 items in multiple-choice format

(2) Discrimination of intonation patterns - 38 items requiring the correct interpretation to be chosen from a choice of two

(3) Reading comprehension - 49 items in modified cloze type format covering two passages with the first letter of the deleted item provided. The items for deletion are non-random.
Time for test 3 : 5 minutes only

(4) Grammatical structure - 47 items for completion with alternatives provided in multiple-choice format.

Tests 1 & 2 are recorded on tape with printed answer sheets.

Part II

Reading speed - an optional test comprising a 'passage of prose of about 700 words into which a number of words, some foreign and some irrelevant English ones, have been inserted. The candidate is asked to underline the irrelevant words'.[1]

---

1. Quoted from the information handout on English Proficiency Test Battery prepared for distribution by The British Council.

3.3.5 The English Language Battery (1970), ELBA, will be the last test to be reviewed and is in use at Edinburgh University (Ingram, Catford, Mackin, May). It includes 7 sub-tests and takes 90 minutes to complete.

Part I

(1) Phoneme recognition - 100 items in multiple-choice format

(2) Intonation - 10 items in multiple-choice format

(3) Stress (Word) - 10 items in multiple-choice format

(4) Listening comprehension - 30 items

This part is recorded on tape and requires answers to be marked in a printed booklet.

Part II

(5) Grammar and structure - 50 items in multiple-choice format Time : 15 minutes

(6) Vocabulary - 50 items in multiple-choice format Time : 15 minutes

(7) Reading Comprehension - 20 items in multiple-choice format covering 4 passages of 80 to 150 words. Time : 20 minutes

3.4 SUMMARY OF REVIEW OF TESTS

The comparatively detailed information provided for both the EPTB and the ELBA is because they will be referred to later in the thesis. These two tests were used to validate the measuring instruments developed during this study.

Both are in extensive use; the ELBA by more than one British university, and the EPTB or Davies[1] test as it is generally known, is used in many countries where the British Council provide help in administering it for various reasons.

In 1965, Pilliner[2] made a comparison between EPTB Version A (1964) and ELBA (1964) and reported a correlation of ·685, significant at 5% level with a sample group of 59.

Our survey has been extremely sketchy since in many cases the tests are confidential, administered by the university for its own purposes. What is of interest for our study is the component called reading comprehension. It generally forms a small part of the battery, and in the British group, is not always isolated for testing but subsumed under grammar and vocabulary. The increase in length of passages where used from the sentence to the paragraph suggests an increasing order of difficulty though this is not always explicitly stated.

If, as we said in the last chapter, reading plays such a crucial role in academic success - and few would seriously contest this - it seems imperative that this aspect of language proficiency be taken more seriously. Moreover the tacit assumption seems to have been made in

---

1. Named after its author Dr. Alan Davies who developed the first version in 1964, and subsequent versions in 1965, 1973, 1976.

2. Pilliner (1965:26) A Comparison of two English Language Proficiency Test Batteries.

many of the tests that four, or five, or eleven short passages represent an adequate sample of the kind of reading students are required to do. What is also assumed is, therefore, that ability to understand the sentence or short passage, or grammatical items, is equivalent to the ability needed to comprehend larger units of continuous prose.

We recognise that in testing, the two broad variables on which success of the entire operation hinges are: (i) the extent to which the task is truly representative and (ii) the extent to which the candidate's performance is truly representative.

Because of the uncertainty associated with all attempts to elicit the latter, much of the attention in testing research has concentrated on the former.

We hope in the course of this thesis to present fairly convincing evidence that far more information on reading performance can be obtained by using the long passage in a testing situation. Not only does it provide more reliable data regarding the candidate's reading ability, but is more representative of the academic reading in terms of length.

We have in the survey stated the type of question form used to elicit performance. With two exceptions which will be briefly discussed, all the remaining tests have relied on the multiple-choice format. This is understandable because with the large number of testees involved, the ease,

objectivity and reliability of scoring must be acknowledged as playing a very important part in solving administrative problems. Also since these items have undergone considerable trial and experimentation in the development stages of the tests, their validity too has been ensured since only those items that have successfully discriminated between more and less proficient foreigners have been retained in the final test version. Unfortunately, we have little confirmation regarding validity of the tests as a whole, or even of the reading comprehension component. Partial scores, we are warned, are not to be taken as representative of the sub-skill they aim to test.

The two tests that use other testing techniques are the EPTB and the UCLA. The former uses a non-random modified cloze, and correlations with tutor's or end-of-course assessment range from ·4 to ·7, depending on different samples. More interesting, the EPTB is the only test to use the intrusive word technique to test reading speed. Unfortunately however, this very important aspect is not taken into account when test scores are evaluated - it remains an optional extra. The UCLA, ESLPE test also uses the cloze technique, but here the use conforms to the version as first proposed by Taylor (1953) and which has been extensively reported and researched. It also uses another technique that appears more congenial to the classroom - the dictation. Both these techniques present examples of an integrative type of testing where overall proficiency is represented.

Both Oller et al (1971:190) and Darnell (1968:40) have reported fairly high correlations between cloze type tests and two of the standardized batteries we have reported. Oller reported a correlation of ·87 between UCLA and his cloze test with a sample of 35 non-native speakers (scored by acceptable word method) and Darnell reported a correlation of ·838 between TOEFL and his test using cloze for a group of 48 NNS.

These latter studies using correlation coefficient to establish validity of their cloze tests have adopted the stand that cloze measures proficiency in English as a second language, or as with EPTB, that it is a valid indicator of reading ability.

3.5 EXPERIMENTS IN READING

By contrast the experiments in reading have questioned basic assumptions made regarding the relationship between language and reading ability. Here too the concern has shifted from the task to the learner's performance as reflecting ability. One of the most striking features is the use of longer texts in many of the studies. With the growing interest in discourse analysis, experimenters are explicitly questioning the validity of extending skills based on short texts to include MVD. In the summary of reading research published yearly by the International Reading Association in the Reading Research Quarterly, a brief comparison over three years will make this changing attitude clear. A survey of the research on 'Experiments in learning' published in the spring issue contains:

1973-1974 Seven studies with one using a long passage

1974-1975 Thirteen studies with one using more than 1000 words

1975-1976 Twenty-six studies with thirteen using more than 1000 words, and two more using 800-1000 words.

The time taken for reading a passage of about 1000 words by native speakers has been fairly stable at 6 minutes in 1973-1974 and ten to fifteen minutes for 1500 words in 1975-1976.

In the past, most experiments in learning used lists of words for various tasks that relied heavily on short term memory. To date, recall studies have investigated the effect of presentation order, stimuli reinforcement, stimuli variation, mode and rate of input, duration of learning and strength of retention among others. Most have required recall following shortly after task completion, but increasingly interest has been expressed not in the verbatim recall but in the changes that occur with delayed recall. They have found that memory tends to be meaning preserving and gist is often remembered long after the actual verbal form has been forgotten. The latter have also used longer texts to see how encoding into long term memory has imposed re-organization on input. The influence of context in MVD and its effects on memory storage are areas now being researched.

3.5.1 Miller and Selfridge (1950, 1966) studied the effects of contextualisation on memory and learning. Verbal context was defined as 'the extent to which the choice of a particular word depends on the words that precede it' (1966:270). This particular definition is understandable in view of the way lists of words were constructed for use in the experiment. Starting from a completely random collection of words taken from the Thorndike-Lorge list they ranged in varying orders of approximation to give quasi-contextualised lists, e.g. subject 1 was presented with one word and asked to suggest a one-word continuation. This second word only was then presented to subject 2 who offered one word more, and so on until the list grew to the required length. Similar lists using two-word, three, four, five and seven-word contexts were constructed, and sequences from ten to fifty words were made into a total of 32 lists.

These were then presented to groups of students to be learnt for recall. Order of presentation was to be preserved but in fact all verbatim recall was scored as correct regardless of sequencing.

The results indicated that for the shortest ten-word list, subjects could recall correctly with as little contextualisation as the second order approximation to English. With the twenty-word list, the third order or three-word context was needed, showing that with greater demands on memory, increasing context was of help in learning.

But with the thirty and fifty-word lists, even partially successful recall was possible only with five and seven-word contexts. The authors concluded by suggesting that 'contextual dependencies extending over five or six words permit positive transfer, and that it is these familiar dependencies, rather than meaning per se, that facilitate learning' (1966:278), implying that ability to chunk sequences made remembering easier. This implied that syntactic dependencies were as important as meaning in recall. The longer the passage, the greater the usefulness of contextual associations extending over long sequences of items. For verbatim recall especially, meaningful material was easy to learn because it provided the short-range associations familiar to subjects of the language.

Moving away from lists to experiments with connected text, we find a wealth of literature concerned both with measurement of reading ability and with measures to improve the quality of learning from reading.

3.5.2 Rothkopf (1972) gave groups of subjects a long passage from 5000 to 15000 words with relatively high factual content. Questions were inserted at varying intervals, say every 750 words, and students in the group reading them were asked to answer them when they occurred in the text. Some wrote answers, others merely read the questions, some received immediate feedback regarding the results of each question, and some didn't. Unlimited time

was given for the entire operation with one prohibition, they were not allowed to turn back to an earlier part. Shortly after completion they got a criterion test to evaluate how much learning had occurred.

Inserted questions for this and similar experiments were regarded as 'adjunct' and represented incidental learning, while those items contained in the criterion test were designated essential information content. The aim of this experiment was to study 'mathemagenic activities' defined by Rothkopf as 'those student activities that are relevant to the achievement of specified instructional objectives in specified situations or places' (1972:324). Originally 40 questions were prepared, each independent in content; of these half served as adjunct and the remainder as criterion questions. The purpose was to see the extent to which inserted questions increased students' attention and improved learning or exerted an inhibitory influence.

Different groups had pre-questions or post-questions, questions relevant to the portion of text they related to, or questions far removed from the relevant part. In addition a control group read the passage uninterrupted and then sat the criterion test. Rothkopf by analyzing the results could study the effect of question-position, question-interval, duration of time needed for different portions of the text and duration needed for different questions. He reported (Ibid:319) that content features appeared to be a more powerful influence than structural features as predictors of learning. Where inserted questions asked

for general information or detail, attention to this aspect was greater and helped in similar items which were included in the criterion test. He suggested that questions were a form of environmental pressure that influenced the underlying processes necessary for learning to take place.

There is at the moment contradictory evidence emerging regarding the influence of question placement (Hesse and Smith:1975) and for those interested they are advised to refer to journals.[1] One of the major drawbacks of these studies seems to be the choice of text. The experimenter is forced to choose passages relating several facts to enable him to compile a large number of independent questions. He then often arbitrarily decides to designate some as adjunct or incidental questions, and others as the main focus of learning. Meaning imposed by the structure and composition of the text is not allowed to weigh in the decision since from each part of the text, one question must be incidental and one must be essential. Because questions and their use have such far-reaching implications in the applied fields, the warning comes that very little is known in fact. Rothkopf, as well as Frase (1972:354) recognise the possibility that organizational properties of the text could alter the processing demands made on the reader. They also imply that if content

---

1. See especially the Journal of Educational Psychology, Journal of Experimental Psychology and Journal of Verbal Learning and Verbal Behaviour.

affects the learnability of a text, it follows that readability indices that take account only of word length, number of syllables, sentence length, must be less reliable than measures that also include the type and number of lexical items.

3.5.3 In a specific attempt to investigate the influence of language organization on ease of understanding, Peltz (1974) tested the effects on reading comprehension of re-patterning passages to approximate the syntactic organization of students' writings. He asked 34 students to write a 1000-word essay on the content of a social studies' unit and, attempt in 250 words, free response answers to four questions. The data thus obtained was analysed syntactically applying Kellogg Hunt's T formula (1966). The patterns mainly of subordinate clauses to a main clause formed the basis of the re-patterning. Eight social studies texts were re-written using the structural patterns identified in terms of clause complexity.

Sixteen cloze tests were then prepared using the eight original and the eight modified texts. In addition multiple-choice tests of five items each were constructed. The latter revealed no difference between the groups but the cloze tests revealed significant differences and was judged to be far more sensitive to language change. Higher scores were obtained from the texts which reflected the students' writing patterns. Comprehension was defined here and in all the literature by the number of items correct

on either the cloze or multiple-choice tests. The author has argued very persuasively that the reader's approach reflects his personal expectations and experience, and the degree to which they match the author's language, determines ease or difficulty level of the text for the reader.

3.5.4 Another factor that can affect reading comprehension and which is being increasingly studied in the applied linguistics field is the purpose for which the reading or learning activity is undertaken. This is partly what Rothkopf tried to show in the experiment reviewed under 3.5.2. Samuels and Dahl (1975) conducted two experiments to investigate the extent to which knowledge of purpose affected reading rate. The first study used fourth grade students and the second used 84 college students as subjects. As we are interested in adult readers we will report the second one only. Four groups were used to test the extent to which reading for specific questions encouraged flexibility. Four articles, each about 820 words long were selected, rated as just below undergraduate level by readability measures. Two of the articles were to be used as practice material, and two to be used in the experiment proper.

The design adopted was as follows:

Group 1 : Passage A(general questions) + Passage B(detailed questions

Group 2 : Passage B (detailed) + Passage A (general)

Group 3 : Passage B (general) + Passage A (detailed)

Group 4 : Passage A (detailed) + Passage B (general)

Each passage was followed by a test of three items in multiple-choice format, and before each test passage, the group did a practice passage with questions so that they knew exactly what to expect. The results showed that for the general questions' passage, the rate was faster at 461 wpm with a mean comprehension score of 2·71; while for the passage accompanied by questions on detail, the rate was slower at 291 wpm and the mean comprehension score was 2·42 out of a possible 3. The authors claimed that the explicitness of the objectives was responsible for the flexibility rating not found in many of the earlier studies.

3.5.5 To reinforce the results which hinge on expectations, an experiment by Thomas and Augstein (1972) provides an interesting contrast. In this study of 60 seventeen-year olds, the authors required the subjects to read 500 lines of an article. Two tests A and B were constructed, each containing 22 items. Of these 8 were multiple-choice items testing for simple knowledge of facts, 7 involved short responses on more complex ideas, and 7 asked for recall of simple but specific information.

Four groups were formed. Groups 1 and 2 did one of the tests before reading and were told that it would be administered again after they had read the passage and, in addition, another similar test would be given. Groups 3 and 4 were told to expect a summary question. The following data will show what each group expected to be required to do and what in fact they actually did.

Group	Expectation		Actual Performance
	Pre-test	Post-reading test	
1	Test A	Test A + Test B	Summary + Test A + Test B
2	Test B	Test B + Test A	Summary + Test B + Test A
3	-	Summary	Test A + Test B + summary
4	-	Summary	Test A + Test B + summary

The results indicated that the groups who had read expecting a summary were better able to answer the objective type tests than groups 1 and 2 who had read for details. Since the experiment was not concerned with speed, subjects took from 70 to over 96 minutes to read the article at their own rate.

All four groups were re-tested one week later without warning and groups 3 and 4 showed better retention of information as compared to 1 and 2, the groups who had read for detail.

The authors suggested that learning was affected by the pattern of reading generated to suit the expected tasks. Although instructional directives do influence the reading, they do not determine the strategies adopted by the reader. The latter were probably self-generated as a result of expectancies. Thomas and Augstein thought that search patterns for larger structural properties could lead to a better understanding of the text and so result in more effective retention; and that assessment procedures could influence the operation of the task if they were known

to the reader prior to the activity.

3.5.6 In another study, Schallert (1976) examined the relationship between depth of processing and context, and its effect on memory organization. This, like the Thomas and Augstein experiment, represents a present trend in reading comprehension literature, where non-verbatim recall of content is evaluated. Schallert used a technique increasingly used over the last few years and which we will also use. She listed the number of propositions or semantic 'ideas' in the passages she selected, and used these to score recall rather than using verbatim measures like word counts, etc. College students were used in her experiments. Three paragraphs were selected, each containing among others, four ideas that were ambiguous permitting two interpretations, one of which was the more obvious and called the 'strong semantic interpretation', and contrasted with the 'weak semantic interpretation' which represented the less obvious meaning. The paragraphs were provided with titles giving either strong or weak meaning. The provision of this context enabled the author to compare between groups to see how they resolved the ambiguous ideas and the extent to which the title stimulus affected both meaning and the amount of information retained.

Subjects were instructed to do one of the following tasks when they read each passage; three tasks for three passages:

a strong semantic task (a) rate passage for ambiguity
or (b) rate passage for learning difficulty
a weak semantic task (a) count the four-letter words
or (b) count the pronouns

Exposure times for the reading tasks were either 20 seconds or 45 seconds, and 5 seconds were given for the response to the task to be recorded.

The variables in Schallert's experiment then were: type of task, context cues and duration. After the reading response, all students regardless of group were given a 60-second simple arithmetic problem to preclude immediate memory effects.

Then they were asked to recall each of the three paragraphs in the order of reading. To help them, the first few words were supplied as prompts. Verbatim recall was not required, and seven minutes were allowed for this part of the experiment. This was then followed by a multiple-choice test of 8 items on each of the passages, again in the same order of reading and of recall. Of the 8 items, two were on general relevant information, two on general irrelevant information, and four on the ambiguous ideas.

The results were then analyzed and Schallert writes that the effects of context were evident on the interpretation given to the ambiguous items in the multiple-choice test. The amount of retention was much greater when the reading task assigned was one of the two strong

semantic tasks. This confirmed the general hypothesis that comprehension and memory are strongly affected at input if deep semantic processing is activated. With greater duration, more information was retained, although interpretation remained as for the 20 second reading. She concluded that more information was retained under the more semantic task instructions than under the less semantic stimulus.

3.5.7 One more experiment conducted in 1941 by Robinson and Hall, not often referred to in the literature is of interest since it is predicative of a present awareness that it is unreliable to generalize reading ability from one subject to another. The experimenters chose texts of 3000 to 4500 words at a time when very few studies used anything like this length. Articles from the same source but on four different areas of content were selected, art, history, geology and fiction. Questions on general information rather than detail were constructed paralleling the order of occurrence in the text. 205 students were evaluated both for rate of reading and comprehension accuracy and the authors found that results across subject matter were unreliable. But the interesting thing reported was their finding that for the first three minutes, most of the readers seemed to react in the same way. But with increasing length, variations in behaviour occurred as the material to be assimilated grew and the pace had to slow down to accommodate more information, and individual strategies

had to be used. Reliability varied from the first three to the third three minutes of reading time, and results indicated that 'rates over short periods of time are not indicative of production over longer intervals of time' (1941:246).

Since one of our hypotheses will be that length is a good discriminator between levels of reading ability of non-native speakers, this finding reported for native speakers appears particularly relevant.

The same authors then constructed a fifth test and compared the reading rates of the upper and lower fifths of another group of 285 students. They reported that the latter group revealed very little flexibility in adapting to the varying levels of difficulty in the material. By comparison, good readers appeared far more sensitive to difficulty level. The inference made by the authors was that 'the ability to benefit by increased rate from present comprehension of reading to later sections of the text must represent a higher level of reading ability' (1941:250).

To briefly recapitulate, our review and choice of experiments in reading has attempted to show:

(i) the influence of context on the learning of longer texts

(ii) the influence of question stimuli on reading rate and comprehension

(iii) the effects of expectations on reading rate and comprehension

(iv) the influence of repatterning language on ease of comprehension.

(v) the effects of the depth of processing and context on memory organization.

(vi) the effects of content and length on reading comprehension.

3.6 EXPERIMENTS DISCRIMINATING BETWEEN GOOD AND POOR READERS.

We will conclude our survey by briefly reviewing some experiments relating to discrimination between different levels of reading ability which was briefly referred to in 3.4.

3.6.1 Weinstein and Rabinovitch (1971) asked whether syntactic structure facilitated learning and recall in good readers. They prepared two lists using nonsense words, and structured one using syntactic markers to resemble English, so that the nonsense elements were meaningfully related in the 'structured' list. Two groups identified as good and poor readers on the basis of standardised tests were asked to learn both lists. Results indicated that poor readers took the same time to learn both lists confirming the belief that they were not paying attention to or taking advantage of the syntactic structuring where available. Good readers on the other hand learnt the structured list in half the time needed to learn the unstructured list. Learning here was ability to recall verbatim.

3.6.2 Perfetti and Goldman (1976) were interested in seeing the extent to which contextual cues, both syntactic and semantic aided the reader in encoding information not only in short term memory but in long term memory as well. They used connected discourse as learning matter. They suggested that children with poor language comprehension suffered from one of two possible problems. Either they were relatively unskilled in encoding the linguistic information in working memory, or alternatively, reading comprehension difficulty was a function of textual organization unique to reading and not reflective of the level of language ability in general. They carried out two experiments testing ability, one with digits and one with connected discourse, MVD. For the latter they used a probe technique that interrupted the reading. The question asked was: Did the amount of information encoded proceed without regard to syntactic cues, or if the probe interrupted the discourse before a syntactic boundary, was only part of the meaning encoded?

A sentence was presented as part of a paragraph, and the reading was interrupted at varying points in the discourse. Sometimes the stoppage occurred at syntactic boundaries, and at other times in the middle of a clause or phrase.

The experimenter then used a probe word to initiate recall, and the subject had to try to remember as many words verbatim as he could following the probe word. An example

cited in the report helps to clarify their technique.

'It had been a beautiful day for rowing. When a thick fog arose, Nick ... [interruption]

Given the probe 'when', subjects found recall easier up to syntactic boundaries especially when the interval between the occurrence of the probe word and the interruption was not too great. There was less success with increasing intervals. Subjects however found it easier to recall elements up to boundary markers, rarely across them. This was true of the good readers suggesting that the encoding was probably aided by syntactic chunking so that as long as the probe occurred at the beginning of the chunk, the entire unit could be recalled. If the probe occurred in the middle or near the end of a syntactic grouping, recall was more difficult. Poor readers however, as designated by standardised tests of reading comprehension, had difficulty with recall regardless of positioning of the probe word or the interval between its occurrence and the interruption.

Although the experiments did not confirm either of the two alternatives, the authors felt that the tendency suggested by their experimental evidence favoured the hypothesis of general language deficiency in poor readers.

3.6.3 Much of current research is aimed at tracing the causes of reading disability, and many of the experiments attempt to investigate the cognitive processes involved in the reading activity. One of the ways of diagnosing

mental ability is by studying cases of failure or weakness. Hence the focus on the bad or poor reader, and his ability to identify units like the syntactic grouping or the word.

Oaken, Wiener and Cromer (1971) and Steiner, Wiener and Cromer (1971) have favoured the theory that reading disability might be a specific deficiency connected with reading input. In beginning readers, word identification in tests is used to discriminate between levels of reading proficiency. The authors argued that groups matched for word identification ability, but classified by teachers, or on other tests, as good and bad readers, must suggest that difficulty arose with understanding of connected prose. They presented the groups with cloze-type tasks in four input modes: good reading on tape, material badly read also presented on tape, good clear printed matter and bad, difficult-to-decipher printed matter. Subjects had to orally suggest a filler for the deleted item which was recorded. If an unacceptable and incorrect response was offered, or no response at all, the experimenter supplied the missing item. This insured that for the next cloze item, the input had been the same for all subjects. Results indicated that those designated 'good readers' gave a significantly greater number of consensually and syntactically correct word-insertions than the 'poor readers'. One explanation put forward was that organization into patterns and units of printed matter was the cause of differences in comprehension. The good reader was better at compensating for faults and deficiencies in the input mode, although

both groups did equally well under good auditory conditions. We will see in the next experiment that this particular finding has varied with replication experiments.

Steiner et al argued that it seems as if poor readers are forced to make a cloze response from a vastly greater word repertoire than the good reader (1971:507). The latter delimits the response possibilities utilising both syntactic (correct class form) and contextual cues (i.e. if the passage is about hunting, the lexical item will be from the same field). The good reader in chunking appropriately, has the benefit of using constraints from a larger unit. The poor reader, by contrast, using immediate constraints, chunks less efficiently because the processing of input in storage is too slow to allow for a sufficiently large input. There is less organization of material into critical groupings and this leads to poor reading comprehension.

Since the authors had matched for word identification and found equal performance with good oral recording, they felt that reading deficiency is a specific type of disability associated with input mode and peculiar to printed text.

3.6.4 Sticht (1972) however using adult readers divided into good and poor groups found different evidence when he used auditory and visual input modes. He found no difference within groups but a significant difference between groups for both sources of input. His investigation therefore favours the view that a general language deficiency

characterizes poor readers and is the common cause for poor performance under both conditions.

3.6.5 To resolve this question of opposing theories, Berger (1975) conducted various experiments to investigate the relationship between reading ability and oral and written language skills with specific attention to literal comprehension and organizational ability. Two groups designated good and poor on the basis of the Iowa Basic Skills Test were presented with two 200 to 250 word stories on tape and in print. Ten questions on each passage were constructed: four testing for main ideas, three for significant detail and three for peripheral detail. The ordering of the questions followed the sequence of information in the passage. Each student heard one passage and read the second. In each case they were asked to re-tell in their own words what they had just heard or read. Answers were recorded and time given for recall was 6 minutes. This was followed by the multiple-choice test. Good readers performed better on both modes of presentation, and both groups did better on the oral mode. Correlations between recall and questions ranged from ·47 (oral mode) to ·56 (written mode) for good readers, and ·67 for both input modes with poor readers.

Berger reported that her evidence favoured general language deficiency as a possible explanation between good and poor readers.

3.7 SUMMARY

This survey has covered two completely different areas chosen from the vast literature available on reading and learning from reading.

The tests of proficiency used with non-native speakers of English were all constructed specifically to be used as entrance qualification for academic work. They comprised a battery of sub-tests aimed at evaluating skills associated with listening, rules of language tested through the grammar, vocabulary and usage items, and occasionally, speaking and writing and reading. By and large, priority was given to the first two skills. But for the entire test, the preparation and construction were concentrated on selection of a representative sampling of items that contributed to the specific skill and the overall evaluation of language proficiency.

The reading experiments on the other hand offered much more varied fare, focussing on the reader's performance under partially controlled conditions but with many variables contributing to the outcome. This is why each of the experiments reported have contributed to our knowledge and understanding without resulting in definitive evidence of any of the hypotheses tested.

The latter studies used both connected discourse and longer texts and tried to base research design on theories of reading. The various factors contributing to comprehension were isolated for study and testing, and

we saw the interest in attempting to explain what cognitive processes were involved and how they might be affected by different stimuli. The different techniques were as widely divergent as the aims. We had verbatim recall of lists of words, recall of meaningful 'ideas' contained in MVD, recognition of correct ideas presented in multiple-choice format, requirement of completion of items deleted from a text and scored in different ways, and the use of summary as a comprehension measure offered by the subject both orally and in writing. The variety and richness of ideas presently being investigated in the field makes for extremely stimulating reading.

The experimental studies reported, apart from two exceptions, represent research in the last five years. This has been deliberate, not because there is so little that is relevant or useful, but because excellent surveys have already been made covering much of the earlier published work[1].

---

1. The annual survey in the spring issue of Reading Research Quarterly is useful reading. For a survey of 1200 reports published between 1960-1970 see John B. Carroll (1971) Learning from Meaningful Verbal Discourse in Educational Media: A Review of the Literature.

CHAPTER FOUR

THEORIES AND MODELS OF READING : A SURVEY

4. GENERAL REMARKS.

In this chapter we develop our interest in the cognitive processes attending reading by presenting a survey of some of the theories often implicit in models of comprehension.

Language comprehension models have been hypothesised for some time now, and over the years they have reflected changes not only in theory but also in the area of concern dominant at the time. As we saw earlier under Section 2.3 the definitions of comprehension suggested that some theorists interpreted reading primarily as conversion of graphic symbols to sound, while others viewed it as a complex cognitive activity. We also have models that attempt to isolate linguistic components that learners have to acquire e.g. knowledge of grammar, affixes, rules concerning word order etc. We view this last as an acquisition model that is more the concern of a language teaching syllabus, and hence not relevant for our purposes.

The purpose of this survey is to review the literature, and select for description a few models that are both representative of the field, and useful in themselves. None however has been found adequate for our purposes, since each tends to concentrate on a

particular aspect of the process. Those that are comprehensive have not been developed in sufficient detail to permit application.

However we will be borrowing many of the ideas represented in these theories and using them to develop our model. Our performance model will attempt to incorporate the various processes from input of printed English to output of some response that will be taken to reflect 'comprehension', while at the same time keeping within the constraints set by the physical limits of memory and processing capacity as far as they are known. The model will be presented in the next chapter.

4.1 EARLIER THEORIES

The present survey begins with one of the earliest theories often referred to as the 'classical one', Geyer (1970).

4.1.1 The 'Classical' Approach

This approach viewed the span of perception as central to the process and many experiments since then have been concerned with trying to determine what it is that the reader perceives. The study of eye movements, span of the field of central vision, length of eye-fixation, frequency of eye-fixations etc. have all yielded very necessary information that have contributed to our appreciation of the complexity of phenomena involved in the physical aspect of reading. Studies have tried to determine

whether fixation is a unitary steady phenomenon or whether there is a slight tendency to move from left to right with roman scripted material, and right to left with Arabic script. To all this measurable data is added introspective reports that complicate this theory. Subjects have stated that the visual image remains for one or two seconds after the stimulus has been cut off either by masking, removal, or sudden change from light to dark. This after-image permits reading to continue from memory which implies a visual memory-trace that, at least under some conditions, exists for a brief period before fading.

Research with beginners has concentrated on letter and word shapes, and other non-verbal signs such as punctuation, phrasal groupings etc., which can act as cues to guide the eye to its fixation target, thereby implying that this movement is amenable to training and control.

Many psychologists have postulated that visual stimulus is encoded auditorily so that there is only one major decoding process for both oral and visual stimuli. This assumes the primacy of speech as a physical fact and partial evidence for this view is suggested by its proponents in the amount of sub-vocalisation that appears to occur in silent reading. Opposing theorists have argued that sub-vocalisation, a measurable phenomenon, is perhaps utilised by the reader as an aid to overcoming textual difficulty rather than as a necessary part of the reading process. They put forward as evidence for its optional use, the case of deaf children who do not exhibit this tendency Conrad (1972).

Efforts to suppress sub-vocalisation without affecting comprehension have been reported in some experiments, denied in others, Murray (1967).

All however agree that there is a spatial variable partially amenable to training which is the focus of the span of attention in vision; and it is during this fixation that whatever is 'perceived' within the field, is reported. This means that the reading process can, if it be so desired, be described by an analysis of the series of eye-fixations as the reader progresses through the printed matter. This view of perception as a unitary phenomena is now generally considered unacceptable and outdated.

4.1.2 The Reaction : New Approaches

In 1954, Woodworth and Schlosberg (1954:507) suggested that the reading act was better described as a 'continuous process in that the perceptual development of meaning goes on steadily. Perhaps one can think of it as a continuous production process, a machine into which the raw material is tossed by the shovelful'. We see here the view that raw material is converted into meaning. What had been accepted up to this time as being the same thing was now being questioned, and the new attitude represented the hypothesis that perhaps sensory and perceptual processes were not identical. And if they were not, perhaps research ought to be aimed at the perceptual side that worked on the sensory stimuli.

The kind of questions now engaging people's minds

were concerned with processes, variation in reading ability, e.g. How was it that different readers read differently given the same sensory stimulus? What were these processes, and how did they deal with the information passing through the system? Could they be measured, even indirectly?

By the early sixties it was generally accepted that there was more to reading than the ability to say the words aloud. Simple decoding was not enough or sufficient to guarantee comprehension, and therefore although the two did often go together they could not be said to be the same thing. One could read aloud without understanding. There was also the feeling that different elements in the visual field were being processed differently.

But many theorists still viewed the activity as a word by word progression. It might be of interest to note what George Spache wrote in his book Toward Better Reading (1963:65), and which I have for convenience divided into four sections or levels.

> [Level 1] 'The reader first recognises words by their form, shape, structural parts or by the implications of the context. Each word calls forth one or several meaning associations which the reader tries out for appropriateness in this contextual setting. He accepts what seems to be the most relevant meaning or associative thought and proceeds to the next word, again choosing an association which seems logically related to the preceding word. [Level 2] Various groups of words form cohesive

> associations as he reads through the elements of the sentence. These groups of ideas coalesce into the stated or implied meaning of the sentence. [Level 3] The meanings of successive sentences may be combined inductively into the main idea of the paragraph. In deriving the main ideas of the paragraph, the reader may recognise cause-effect, question-answer, hypothesis-proof or other relationships which contribute to the generalization. Or these sentence-meanings may form the basis of original deductions, such as implications or unstated conclusions, or ideas associated with but tangential to the main idea of the paragraph. [Level 4] The reader may go beyond simple comprehension of the literal, implied or tangential meanings to evaluation of the ideas offered. He may question their authenticity, deny their implications, or reject the bias or prejudice present. He may be moved to consult other sources for verification, to check the author's background, to compare the author's value judgements. Finally the reader may utilise the author's ideas or viewpoint in a creative treatment of the same topic basing his own ideas upon those he has read, or refuting them by proper logic or proof'.

This description or rather explanation made good sense to the teacher who would view the teaching of reading as gradually becoming more complex. But from the theoretical

standpoint, the linear progression of input is very prominent, and this is reflective of the times. Spache writing a year later summarised his views saying, 'thus, in its simplest form, reading may be considered a series of word perceptions', (1964:12). The four-level activity had been reduced to the basic element with intra-sentential elements of Level 1 combining successively to give us sentence meaning in Level 2. Levels 3 and 4 which incorporated reasoning appear to have become peripheral.

Ronald Carver, as recently as 1974, was arguing that it was advisable not to confuse reading with reasoning since the latter is largely a matter of intelligence function. He in fact suggested that Spache's Levels 1 and 2 had to do with the reading process since these were essential, and later level reasoning could not function without them. One sees here that the split between those who view reading as comprising two parts has not been patched up. Many still conform to the view that the uses of reading are different from the nature of reading, and we would agree with this in principle. However we do not support the arbitrary dichotomy introduced between Levels 2 and 3 of Spache. Paragraph meaning and abstraction of author's intention and evaluation are all part of 'reading' as we view the process, and are all involved in meaning. We might perhaps exclude the last two sentences of Spache's quotation since this is how the subject uses the results of reading, but we would argue that everything else was inextricably involved in the activity.

To return to our brief search of the literature, by the late sixties, linguistic theory was beginning to make an impact on theories of reading, and analysis of textual matter was revived as an area of interest.

4.1.3 Contributions of Linguistic Theory

David W. Reed in his article 'Linguistic Forms and the Process of Reading' (1970a:19) states that in order

> 'to discuss reading from a linguistic point of view, it is important to distinguish between the elementary aspects of reading (termed here "the process of reading") and what reading specialists are accustomed to think of as more advanced aspects of the same subject (which may be termed "the uses of reading")'. He expanded this by saying 'the process of reading was defined above, by implication, as the identification of linguistic forms through viewing the graphic symbols by which they are conventionally represented in a given language'. What was needed was a more precise understanding of the nature of linguistic forms and the writing system as represented in English. He defined a linguistic form as 'a linking of a unit of meaning to a physical representation in terms of a conventional system such as speech or writing' (1970b:225).

The following diagram makes the relationship clear.

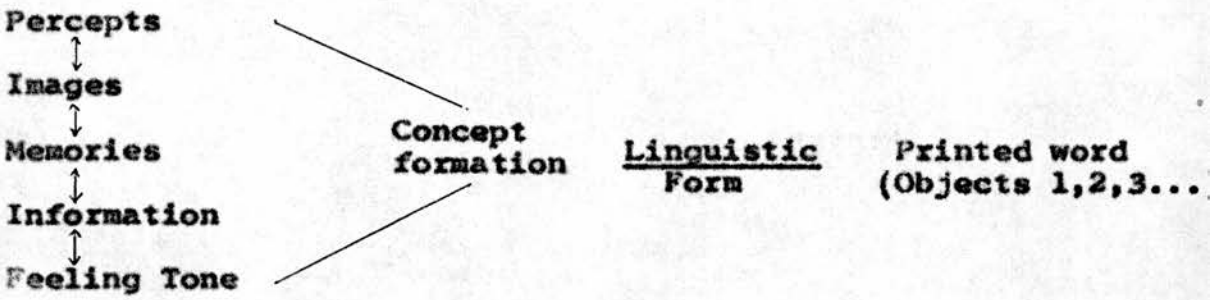


Fig. 1. THE LINGUISTIC FORM

[Referred to by Singer (1970:152) and titled : A sub-system of related elements mobilized as a conceptualized response to a printed word stimulus that represents a class of objects].

For Reed, the linguistic form was an element of deep structure and one of the best ways of proving comprehension was by reading aloud since faulty intonation would indicate failure to understand meaning, which in turn would be directly attributable to the 'failure to identify linguistic forms' (1970a:23). This term he borrowed from Bloomfield (1933) for whom it meant a grammatical unit of a language, and Reed felt that the idea of a 'concept' as distinct both from meaning and from the concrete symbols of speech and writing, would be a useful tool for research into the process of reading.

But none of the views presented so far could be regarded as theories to be applied to language and reading.

In 1966 we have a start.

4.2 THE GEYER MODEL (1966)

John Geyer presented his model which he claimed was a heuristic model aimed at providing for the sensory input which is sequentially ordered. He was concerned with trying to offer psychological explanations for physical phenomena which appeared to have been experimentally established.

The model begins with an input at the rate of 8 msec. a letter space during a fixation which lasts for about 250 msec. Immediately, organizational and associational processes are activated and meaning developed. The latter could well be after input has been organized into a word unit for processing. This system provides for various transformations to occur at each stage in the system.

This organization, Geyer calls 'coding'. From here, the unit moves to a storage device whose function is to provide time for processes to do their work. This store acts as a buffer separating sensory phase from response phase, and Geyer hypothesised that the time limit of this buffer storage was about one second. This fitted very neatly with the visual memory-trace reported by subjects after the visual stimulus had been removed. The store also allowed the unit to grow in size. Once some kind of covert verbal response had been made, the data was transferred to a second storage system. The latter acted as another buffer, this time between the 'recognitional and motor phases'. Geyer suggested that this second storage

could retain units for as long as four to five seconds if needed, and was probably akin to what psychologists have called short term memory or STM. He did not state whether images stored were in visual or acoustic or some other form. We will discuss this model later, but the details follow:

System	Program
Visual Input	-> sequential input at 8 msec per element
Coding	-> conversion to icon, including higher order organization for
Iconic storage	-> icons independently stored for about one second
Internal response	-> verbal responses at about 250 msec per unit
Verbal storage	-> storage limit of several seconds - indefinite storage possible through rehearsal - akin to STM
Verbal Output	-> processing rate in oral reading about 300 msec per unit.

Fig. 2 BLOCK DIAGRAM OF GEYER MODEL OF ORAL READING

The components comprise: (i) sensory input (scanning system).

A major hypothesis in the model is that visual input is sequentially ordered within a fixational pause. This means that one of the major functions of the eye-movement is to ensure fixation points that permit fine discrimination. Other contents in the visual field may be simultaneously processed but through other systems. [What these might be is not elaborated].

(ii) Sensory organizational systems: (a) On scanning i.e. after input, the development of meaning begins and may continue well after the data processing is completed. But the model makes provision for the possibility that certain changes may occur to the material before being passed on to iconic storage, e.g. letters may be grouped to form phonemes, or even words. (b) The iconic store acts as buffer between sensory and response stages to allow processing of data that need different times for completion to take place.
(iii) Internal response system: this covers the whole gamut of complex and varied operations that are needed to convert data to some form of meaning. Geyer suggests that the time taken is similar to that needed for a fixation pause, about 250 msec. This explains the comparative stability of the eye-fixation pause despite age or training - it is the time needed for processing to occur.
(iv) Secondary (internal response) storage system: this now receives the data and can store it if necessary for several seconds. This short term memory is subject to decay, interference and displacement.
(v) Verbal output: this permits the response necessary in a model of oral reading.
Because of the concern with oral reading, the attention paid is to observable phenomena of input and output rather than to the cognitive aspects of the activity itself.

However some extremely useful hypotheses have been presented for consideration. The stability of eye fixations is known, but to explain it in terms of processing

time needed, and the suggestion that different parts of the processing may need differential times is extremely interesting. This means, and much of the experimental work in eye-movement bears this out, that input from four fixations can be held in iconic store for a second allowing the formation of smaller elements into larger units.

It is reported that good readers can take in around 20 letters located within an oval field, equal to three rows of print of about 2 inches across and 1 inch down in the middle. This is about three to four words and there is likely to be overlap between fixations, Gough (1972). Of these 20 or so letters, only three are in sharp focus, Samuels (1970). Good readers need about four fixations to a line compared to about two per word for poor readers. Of the 250 msec. needed to fixate, about 100 msec. are taken up with refining the focus mechanism.

One basic advantage is that the good reader can control rate of input to allow processing to occur so that nothing is lost. If input came faster than could be processed, some of the data would be lost since it can only be stored for one second in iconic buffer. This allows all the input to be output for oral reading if there is careful balancing of rate of input. This allows a smooth even rate of reading aloud. Experiments with EVS, the Eye-Voice Span, indicate that the eye is normally 4 to 5 words, or about one second ahead of the voice, to allow the necessary processing to occur, Schlesinger (1968).

Geyer suggests that perhaps the poor reader is unable to balance rate of input and output which explains the halting, jerky reading, and the frequent regressions or looking back. If the input is too slow, higher units cannot be formed and smooth vocalization cannot be effected.

The major criticism as I see it, is essentially the lack of detail concerning the process. Also it is implicit that if input is sequentially ordered, the processing too is linearly oriented, and more important, all input is processed. How then can we account for reading mistakes, or of such common experiences as 'I looked for it, but didn't see it', and 'It's right in front of you. Can't you see it?'. The choice of verbs in English is itself an intuitive recognition that there is more to seeing and perceiving than the presence of a stimulus and a seeing mechanism.

Another major criticism: Where is meaning to be assigned, and how? What are the major mechanisms involved and where do they operate? We need to remember far more than the time limit capacity of a STM store permits. Where is long term memory? How does one's knowledge of the language and the world affect the decoding process? Too much is omitted for this to be a working model. Geyer has recognised the need for the various components we have mentioned to be included, and offers an accompanying generalised model that includes interpretation and comprehension but without specifying how it operates.

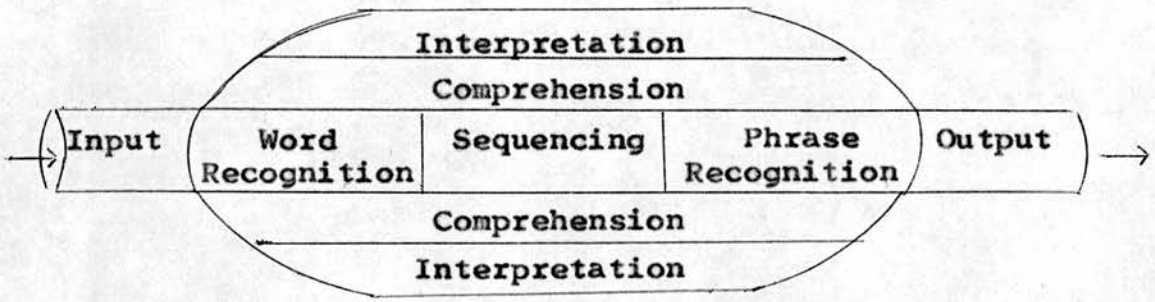


Fig.3. GEYER: GENERALIZED MODEL OF READING

4.3 THE RUDDELL MODEL: A SYSTEMS OF COMMUNICATION MODEL (1969)

A far more interesting, though relatively complex, model is that presented by Robert Ruddell called a 'Systems of Communication Model'. This has been obviously influenced both by communication theory, and later developments in linguistic theory that have tried to accommodate a semantic component. We have here the notion of surface and deep structure, the latter being interpreted in terms of meaning.

Unlike the Geyer model it makes very little attempt to explain what or how much of input is being processed at any given time, but concentrates on the process of comprehension. In terms of the model, the latter depends on what Ruddell calls the relational meaning dealing with structural relationships in sentences, and on lexical meaning which tries to explain semantic considerations of the material read. It also tries to account for non-linguistic information including our knowledge of the world, our experiences of and with things, and various cognitive

strategies language-users adopt and the innumerable variables like interest, purpose and motivation, that have increasingly come to be considered as important as the text. See Fig. 4 on page 78.

What I find of particular interest is the progression from one model to another. Ruddell has tried to include all that was missing from the Geyer model, and in this, represents the enormous advantages that result when a theory has been formulated, even sketchily. This permits a re-definition of the problem and a step forward in knowledge.

The accompanying diagram shows that the model is a far more comprehensive one and includes both spoken and written communication as input, supporting the hypothesis that both deep structure and transformational rules have psychological reality. The model as presented is a little untidy, and I shall attempt to clarify the working by tracing the journey of printed matter from input to meaning which is stored in long term memory.

(i) Input: Spoken and printed matter is input and is either broken down into, or built up into phonemes and graphemes. We have no information regarding the unit of input which could be bits of information, letters, syllables, words or larger groups for written communication. There is a direct link between the two forms, spoken and written, which allows graphic matter to be converted to sound groups if this were needed, Beginners might use both modes of input, but advanced readers can by-pass the phonemic transformation.

(ii) The grapheme now refers to various morphographemic rules, and then enters the cognitive system. This appears to be implied in the model. The surface structure is input into the central processing unit in presumably either visual or auditory mode, but this remains unclear.

(iii) Short term memory is activated as morphemic rules begin their operations. The data remains here briefly to allow syntactic rules to function. We assume that the work is mainly one of 'chunking' smaller units into larger ones. Ruddell refers briefly to experimental evidence that short term memory has a storage capacity for 5 to 6 discrete items, not more. These items may remain for four to five seconds duration but the size of the individual items may vary in size. The larger the unit, the more the data that can remain in short term memory and be chunked. From the diagram, we see that contextual cues can help in the syntactic chunking.

(iv) The chunked material is now subjected to transformational and re-write rules of the kind associated with TG grammars.

(v) But now we have a departure from TG theory, and Ruddell includes a semantic component that dominates the stage. The deep structure [which is represented by roughly the right half of the diagram] must now be analyzed, and semantic interpretation or projection rules are now activated to begin work on the input.

(vi) Long term memory is now activated since rules of grammar of the language are lodged here. These rules work on both the structural and semantic markers present, aided by previous knowledge of the discourse.

(vii) The referrent can now be established, and this is in terms of the pragmatic link outside in the real world.

(viii) The kind of meaning to be assigned must now be determined. Ruddell has suggested that there are three kinds of meaning and all are important : 'denotative', in which contextual constraints can help to resolve uncertainties; 'conotative', which determines that words of high associative value are more easily understood; and 'non-linguistic', which probably refers to the constraints imposed on meaning by non-verbal matter. This could be pragmatic and cultural factors that would delimit the range of possible alternatives in determining meaning.

This part of the model is vaguely defined and Ruddell suggests that perhaps there is a store here of semantic features associated with lexical meaning which will enable the analysis of sentences like:

The boy struck the girl.

This could be rewritten as:

The girl was struck by the boy.

It could also be reversed to give:

The boy was struck by the girl.

But if one had a sentence like:

The boy picked the apple.

Whereas, it could be rewritten as:

The apple was picked by the boy

it could not be reversed to give:

*The boy was picked by the apple.

Semantic rules would have to resolve these differences so that lexical items could be stored here in terms of features of the kind associated with componential analysis.

(ix) The projection rules help in integrating the correct reading of the lexical items with the 'grammatical relationships as indicated by the deep structure to derive the semantic characterizations of sentence constituents' (1969:69-70).

(x) The results of all this analysis is stored in long term memory along with the analysis of the discourse up to this time.

(xi) The model ensures continuous feedback from long term memory so that on-going discourse is affected by what has gone before, although it is not clear whether it can be affected by later input.

Not very clear is how the context feedback that helps to verify information being decoded, actually works. Nor is it clear where the effect of the cognitive strategies and affective mobilisers is felt. It remains outside in a kind of limbo.

Ruddell rightly argues that the logic of postulating a short and long term memory is due to the known temporal capacity of short term memory. Miller (1962) is cited with his experiments which showed that self-embeddings of the kind

'The rat that the cat that the dog worried killed ate the malt'

take far longer to analyse despite the fact that in meaning they are identical to sentences like: 'This is the dog that worried the cat that killed the rat that ate the malt'. The demand on short term is too heavy and long term memory is needed for processing.

Ruddell has also taken account of different surface structures that may be recognized as having the same meaning, given time, by associating deep structure with meaning. A further corollary is that, after surface structures have been analysed for meaning, what is retained in long term memory may have little regard for the original format. Here the work of Mehler (1963) is cited to support the argument that complex structures are often recalled in simpler form.

Ruddell's model includes memory and TG rules etc. but he hasn't clarified what is to be interpreted as components, and what as processes. One assumes that all the rectilinear boxes are components, but nowhere is it specified. There are more questions raised than are answered, but the very fact that it stimulates

one to raise these problems is a big mark in its favour.

His model reflects the impact of linguistic theories and ideas, and in spite of its untidiness is a thought-provoking model. It is far more comprehensive than the Geyer model and we will be using many of the ideas suggested here, although organized differently.

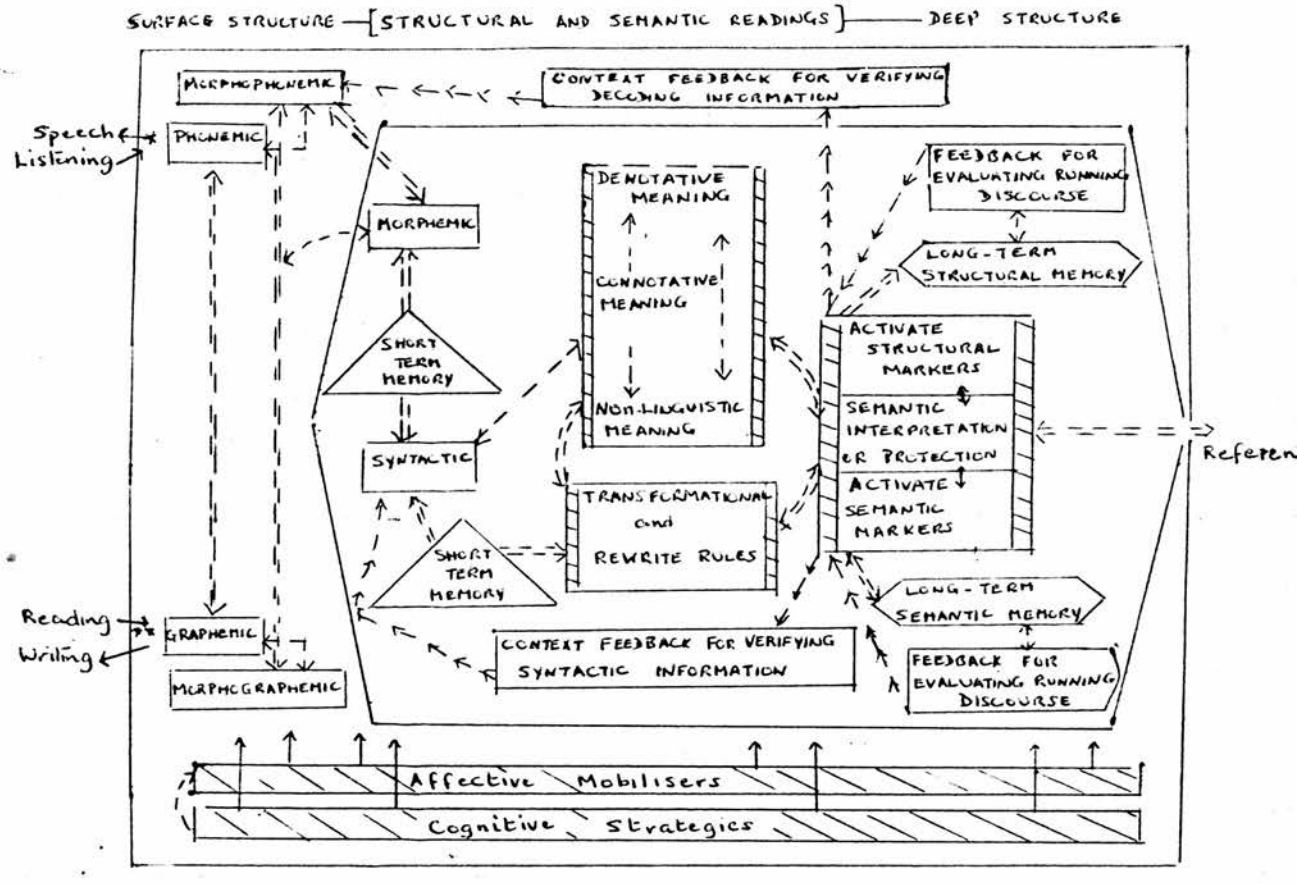


Fig. 4 RUDDELL: A SYSTEMS OF COMMUNICATION MODEL (1969)

Key : * Auditory Perception

** Visual Discrimination

——— Primary Communication Line

– – – Secondary Communication Line

← ← Feedback Line

4.4 THE GOODMAN MODEL: A PSYCHOLINGUISTIC GUESSING GAME (1969)

At about the same time as Ruddell, Goodman proposed his model which has become fairly well known and is often referred to. It brings to reading and the cognitive processing of printed matter, a new perspective, and because it approaches the whole decoding question from a different angle, well deserves consideration. In 1970 an updated version of the theory in the form of his flow chart was presented by William Gephart, and is attached here for interest. (See Fig. 5.)

One of the main arguments put forward is that reading is not a precise process relying on 'sequential perception and identification of letters, words, spelling patterns and larger language units' (1970:259). It is instead a selective process involving 'partial use of minimal language cues selected from perceptual input on the basis of the reader's expectations. As this partial information is processed, tentative decisions are made to be confirmed, rejected, or refined as reading progresses. More simply stated, reading is a psycholinguistic guessing game' (1970:260).

Explaining the working of his model, Goodman suggests the following steps though he says they do not necessarily occur sequentially as presented here for convenience. His explanation helps clarify the flow chart presented in Fig. 5.

(i) The reader scans from left to right along the line and down the page, line by line.

(ii) Eye fixation occurs to permit focus.

(iii) The reader selects graphic cues constrained by 'prior choices, his language knowledge, his cognitive styles, and strategies he has learned' (1970:269).

(iv) A perceptual image is formed based on what he sees and what he expected to see.

(v) A memory search begins for related syntactic, semantic and phonological cues which may lead to re-selection of cues and reformation of perceptual image.

(vi) At this point he makes a guess or tentative choice in keeping with graphic cues so far. Semantic analysis leads to partial decoding which is stored in short term memory as he proceeds.

(vii) If no guess is possible, the reader recalls the perceptual image or looks for more graphic cues.

(viii) If he has guessed (stage vi), he tests it 'for semantic and grammatical acceptability' in terms of the earlier choices.

(ix) If the choice is unacceptable, the eye re-gresses to locate cause of inconsistency. If he cannot solve the problem he looks ahead for more information.

(x) If the guess is acceptable, the meaning is assimilated with prior meaning and any adjustments necessary are made. 'Expectations are formed about input and meaning that lies ahead' (1970:270).

(xi) Then the cycle continues.

Goodman adds that throughout the process there is constant use of long and short term memory.

He suggests that precise identification is unnecessary and the average reader can decode graphic input direct for meaning and then re-encode meaning for reading aloud. But unlike Geyer, he explicitly states that in the activity, the reader uses three kinds of information : graphic cues, knowledge of syntactic and knowledge of semantic information. All three help him to minimise his need for visual cues. 'Redundancy and sequential constraints in language, which the reader reacts to, make this prediction possible' (1970:266). And Goodman very sensibly makes the point that the reader uses cues from both the central area and the peripheral field of vision.

It is worthwhile reminding ourselves that all the theories are concerned with a reader who is decoding in his first language. The ability to utilise cues from different sources is possible if the readers 'are reading material which is fully formed language' (Ibid).

We find a number of the ideas presented intuitively very appealing but the model provides little or no information regarding the organization of the various components in long term memory that must be consulted for meaning to occur. Looking at the flow chart (which incidentally is not really a flow chart at all) there are a number of questions one can ask. Is Goodman demanding more from short term memory than experiments suggest it can perform? If all input

that is processed is automatically stored in long term memory, how does forgetting occur? Would this not be inefficient use of long term capacity? What does he mean by the word 'cue'? Surely one of the functions of a syntactic 'cue' is to signal that what follows is semantically important, or, that it is important in itself. If there are three kinds of cues, a precise definition of terms would be useful. The flow chart using symbols taken from programming needs explaining, and the use of unidirectional arrows whereas necessary in computer language, appears less suitable when used to represent cognitive processing.

The model is far too vague, and it seems to me that Goodman has blurred an important distinction namely the role of cognitive strategies in the reading process. Process and strategy are surely distinct factors. In the context of an individual reader, one recognises that strategies will determine how input is selected and processed for meaning; but it seems important in an abstract model of the process, that the two be recognised as separate. Surely Ruddell's distinction between process on the one hand, and strategies and affective mobilisers which have to do with purpose and motivation on the other, is a useful one.

However, once again, we will be borrowing some of Goodman's ideas for our synthesised model.

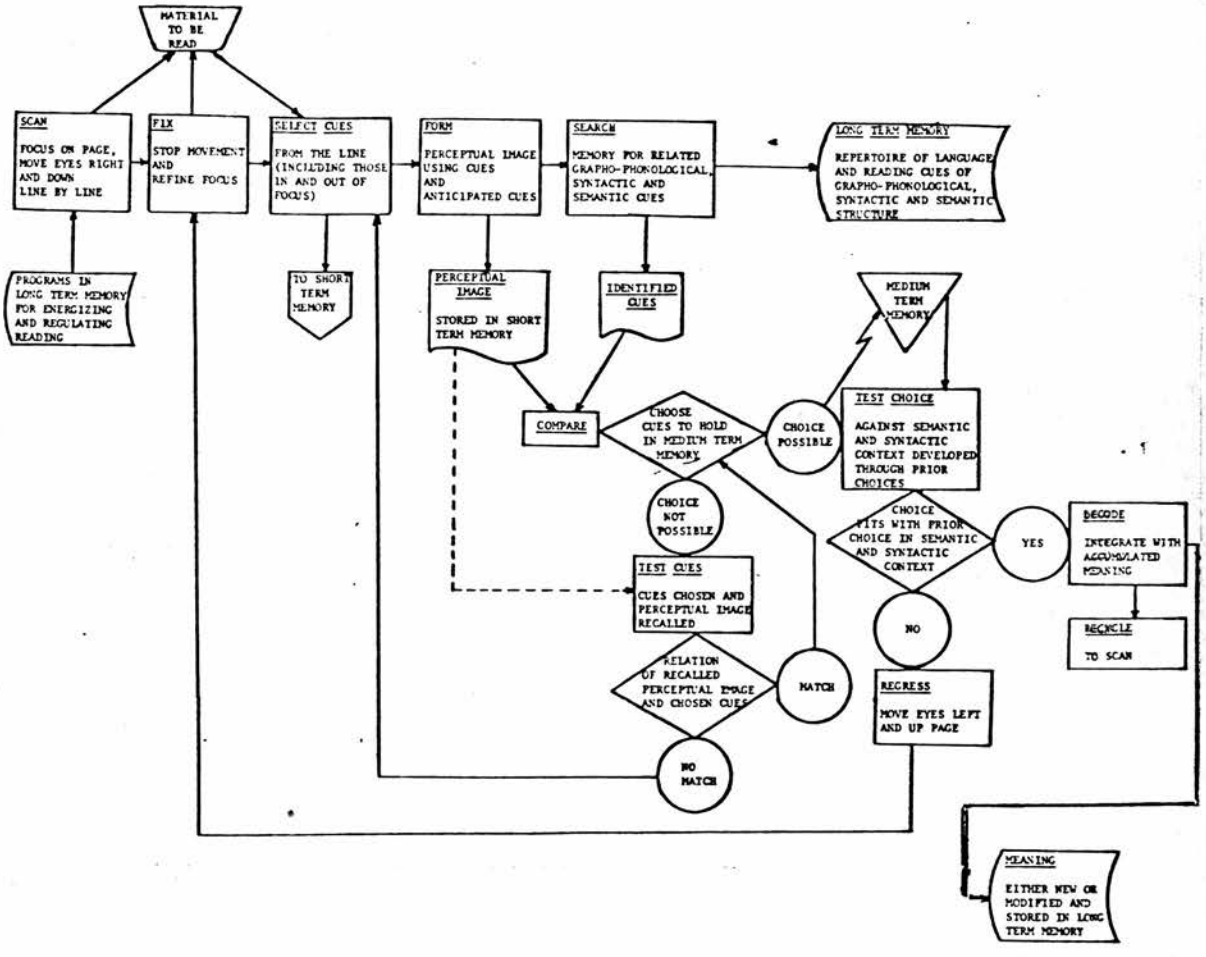
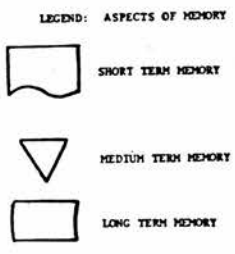


Fig. 5 GOODMAN: A FLOW CHART MODEL OF READING PROCESS (1970)



4.5 VENEZKY & CALFEE : THE READING COMPETENCY MODEL (1970)

Lastly I want to briefly present a model by Venezky and Calfee who concentrate on an area largely ignored by Goodman, that is, the components that must exist in long term memory to enable comprehension to occur. This model is not in any sense representative of the process but presents only some components and their interrelationships that contribute to the process.

The components are outlined below along with the abbreviations used by the authors.

IKS Integrated Knowledge Store : this store contains the most stable knowledge the reader has, e.g. how to read, strategies, sentence types, knowledge of the world etc.

TKS Temporary Knowledge Store : this store contains information of what is currently being read, obtained from the Scratch Pad Store which forms part of TKS. Here are stored details of content, stylistic features, word detail, colour, print etc. of the text. Transfer may occur from here to IKS for some of the information. However transfer from IKS to TKS will depend on need for retrieval and the frequency of this retrieval function.

SPS Scratch Pad Store : this contains that portion of TKS which is necessary for analyzing what is currently input, probably an area where word components are organized and sentence components

sorted out. Information from here is usually limited and decays rapidly unless passed into TKS or IKS. Inter-sentential organization is not performed here since the limits of this organizational ability are generally sentence-bound.

Word Stores:

AWS Associative Word Store : contained here are the most frequently used words along with syntactic and semantic features, and pronunciation features. Access is very rapid to features like word-beginnings, length etc.

LFS Low Frequency Store : words not in AWS, are stored here, and retrievability will depend on usage and how it has been stored, that is, is it accompanied by graphemic cues, phonemic cues or what.

Most of the entries in the word stores will have connections to articulation instructions though some may be very weak.

ILS Integrated Letter Store : similar to IKS, this store contains stable information on letters and letter expectancies [which probably refer to morphographemic rules]. This store indicates to the forward scanner that a particular group is strange or frequent etc.

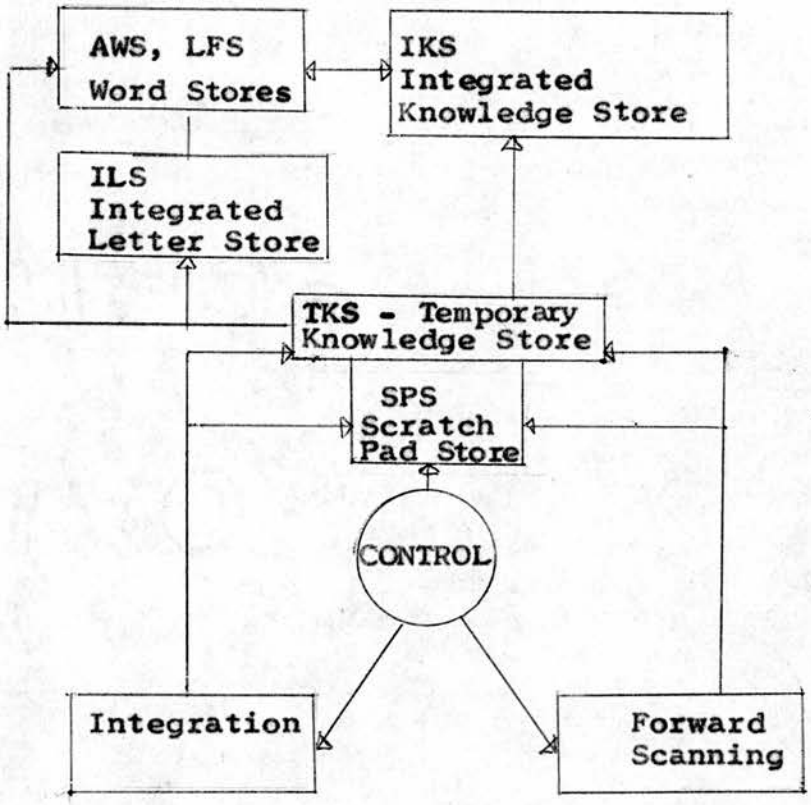


Fig. 6. VENEZKY AND CALFEE : SCHEMATIC OF READING MODEL (1970)

The theory works in the following way. Scanning is central to the reading process and involves a balance between integration and forward scanning, the former dependent on the latter. You will recall that Geyer also made this point of maintaining a proper balance between input and processing. If the material requires careful attention, forward scanning is restrained and the balance is to be determined by the purpose and motivation of the reader, i.e. input can be controlled providing for flexibility in rate of reading. Venezky and Calfee make the same suggestion that Geyer proposed, that perhaps poor readers are not able to shift the balance between the two processes.

Next, the forward scanning determines the largest manageable unit - LMU - which may coincide with phrasal units, punctuation or other features. This is a concept we will be borrowing. Each of the items scanned, letters, letter-units, words, word-units, phrases etc., are now tagged for degree of attention and integration. Integration involves the identification of the tagged items, relating them to previous information and predicting what should be scanned for next. [We see Goodman's influence here].

The authors seem to give to scanning a function of "search" - searching for the LMU which can then be integrated by reference to IKS, Word stores and ILS. The model consists of three processes, forward scanning, integration and finally comprehension, but all three are highly interdependent.

As we can see the theory makes little attempt to relate the model to perceptual processing or even to take account of known facts regarding memory and attention capacity. But it does try to define components the authors feel must form part of the cognitive organization-complex, and suggest an extremely useful device, the TKS, the temporary knowledge store where all information regarding the textual input is collated. Goodman by contrast had all information sent to long term memory which seemed to us an inefficient system.

Another invaluable concept is the explicit component, control. It seems essential that the ability to monitor one's reading activity should be accounted for in some way in a model.

Discussing the model some years later, Venezky took account of recent experiments in using oral miscues to illumine the process, and added, 'For a given reader, the distance to be traversed is probably a function both of reading habit and of the immediate textual situation, that is, how much of the text lying immediately ahead he thinks he can recognize in one fixation, judging from the syntactic and semantic form of what he has read so far and the vague forms of spaces that he observes in his peripheral vision ahead of his fixation area' (1976:11).

He recognized that fixation must occur for a long enough time to permit perception and processing of the intake to begin. And under the growing volume of literature on comprehension as the reduction of uncertainty, he stated that the stimulus variables were word shapes plus context. To this could be added expectancies of what words could occur at a given point in the sentence, based upon knowledge of specific sentence context, prior experience with the material being read, and how familiar the reader was with the words in actual use.

Experiments with oral reading reported by Goodman (1969) and Rose-Marie Weber (1970) provide evidence that 'beginning readers use their knowledge of grammar to narrow down the words that compete for a given sentence slot' (Ibid:162). Correction of errors bear out the greater degree of influence exerted by meaning as compared to the attention paid to graphic signals alone, Kolers (1970). Good readers especially ensured that meaning if not the form

was preserved, whereas poorer readers made fewer corrections and many more misreadings of text thereby confirming the Geyer and Venezky/Calfee hypothesis that they were less able to preserve a balance between integration and forward scanning.

4.6 IMPLICATIONS FOR THE MODEL : PHYSIOLOGICAL AND PSYCHOLOGICAL EVIDENCE.

From our review of the models, it is evident that the reading model presented in the next Chapter will have to take account of a number of findings based on experimental data.

4.6.1 Sensory Phenomena

We know that the eye fixates for about 250 msec. every time it focuses on print thereby allowing for a maximum of four fixations a second. Of this time, the first 100 msec. are reported to be spent on refining the focussing mechanism. The area of vision highlighted is about three letter spaces in sharp focus Samuels (1970:278), with a larger peripheral area, oval in shape and covering about twenty letter-spaces, Gough (1972). However we do not know if input is linear with everything in the field of vision necessarily imprinted on the retina, or if vision is selective with a slight movement involved during fixation.

But we do know from experiments reported by Hochberg (1970:87) and Levin and Jones (1968) that letter spacing is taken into account and can affect reading

performance. Using texts with inter-word spaces eliminated by firstly a letter symbol, and then with an asterisk, Levin and Jones reported that loss of speed was noticeable among the good readers. This was taken to indicate that good readers used spacing to direct eye-fixation. The absence of effect on the poor reader was taken as an indication that word by word reading was being done so that the span of input was either equal to or less than word-length. But the variation in performance was taken to support the belief that rate and amount of input were amenable to control and training.

We also saw from experiments reported in Section 3.5 that contextualisation influences reading and learning, that is, there is a lower threshold of recognition for contextualised matter. But there is more to reading than what literally meets the eye. Reading speeds instead of being stable for NS or NNS are influenced by the rate at which the individual can process input. The reason for reduced rate must lie in the cognitive processing involved and in what happens to the printed matter after input.

4.6.2 Limitations of Attention and Processing Capacity

Miller (1966) reports that among other things our processing capacity is affected by at least three kinds of limitations that need to be kept in mind. In his article, 'The Magical Number Seven, Plus or Minus Two : some limits on our capacity', he reports that experimental evidence suggests that:

(i) Our span of absolute judgement is around seven for all unidimensional stimulus.

The implications here are two-fold. If one could make decisions based on relative judgements, more information could be processed. And if one could judge discriminating details from more than one dimension, perhaps three or four, the amount of information that could be input and so processed, could easily be quadrupelled at one time.

(ii) The second limitation is imposed by the span of attention.

At one glance, the limit of different items seen is around six. Little is said about the size of the item so that once again the possibility exists of varying the total amount attended to at one time. What about attention to non-visual phenomena? We know that mental operations need attention capacity, and duration can continue unless affected by fatigue or interference, but the number of discrete items appears to be limited.

(iii) The third kind of limitation is the span of immediate memory, and this, too, directly affects the reading process.

The span as reported is seven, plus or minus two. Here again the size of the item that can be held in store can vary, so that the amount of information contained in one item would increase if

we could 'chunk' or organise our information into larger units (Ibid). 'By organising the stimulus input simultaneously into several dimensions and successively into a sequence of chunks we manage to break (or at least stretch) this informational bottleneck' Miller (1966:265). This is how we overcome the severe limitations that span of absolute judgement and span of immediate memory impose on the amount of information we are able to receive, process and mmember for immediate recall and for reading.

4.6.3 Organization of Memory.

One of the best theories of memory and attention was put forward by D.A. Norman (1968) in which he included perception, attention and storage facilities. And although many of our ideas are similar, the reorganization of different components is affected by rather different approaches. Norman felt that short term and long term memory were two properties of the same store, he also felt that all input was automatically processed until input was matched with items of meaning in the store. Then, and only then did attention and selection begin to determine which items needed further processing. We hope to show that perhaps our model is intuitively a better alternative explanation for non-native users. But to briefly reiterate experimental findings, Geyer reported in his model that immediate memory has a limit of around one second before the trace begins to fade or is lost due to displacement by influx of new items.

He called this immediate memory the 'iconic buffer'. Subjects' ability to continue to read from memory when the stimulus is suddenly cut off supports this view of the limitations in capacity attributed to immediate memory. Four to five words taking one second to read would be held here allowing reading to continue. This is what is referred to in the literature as the eye-voice span, EVS. By allowing the eye to be four to five words ahead, the reader permits processing of the input to occur so that when the words are read aloud, they are meaningfully interpreted. The EVS also permits smooth reading aloud to occur. Although the eyes could intake at a faster rate, the speed is controlled by the limits imposed by processing time working on the input, and will adjust to reflect the proficiency or inability of the reader.

There is another kind of memory often used in experiments, and called by different names. This is primary or short term memory, PM or STM for short, where information can be retained for a slightly longer period, about four to five seconds. Recall experiments usually utilise storage capacities and organizational tendencies of STM, and evidence of input must occur immediately or soon after learning is completed otherwise forgetting will occur.

Once again the implications for reading are of prime importance. If input of three to four fixations can be sent to STM, the four to five words can now be

'chunked' into a larger unit so that higher level syntactic decoding can take place. Experiments have shown that with average readers, decoding does occur in groups or chunks rather than with single words, and that chunks tend to correspond with syntactic units of a sentence, Schlesinger (1968:42). This is one area where Norman (1968) disagrees with Miller and others. He favours the view that decoding is immediate with a delay of at most one word because in his theory, decoding is automatic. We agree with Miller and Schlesinger since we feel that most of the evidence to date supports the chunking theory.

As input increases, larger units of verbal input cannot indefinitely continue to enter STM with its limited durability. If we wish to remember information after a time interval, we must ensure that we provide opportunities for rehearsal keeping the items active, or re-read the same visual stimulus, or vocalise the information by repetition. This is the only way we can prevent them from being lost due to displacement from STM. Alternatively, we could ensure their survival by transfer to a memory store with more permanent power, and from where items could be retrieved on demand. This brings us to one of the more exciting new advances in chemical technology which provided us with very interesting data directly affecting our theory, and making one of Norman's inferences untenable (Ibid).

We have known for a long time that memory was tied up with things remembered and things forgotten. Questions

that have engaged the attention of educationists and psychologists have been concerned with why some things are remembered, and others forgotten; why some people are better able to recall things than others; why we can't remember everything we want to remember. Although we haven't been given the answers to these questions, we are told by David Krech that 'the physical basis of any memory, whatever else it may be, involves either the production of new proteins, the release of differentiated molecules of ribonucleic acids (RNA's) or the induction of higher enzymatic activity levels in the brain. In a word for every separate memory in the mind we will eventually find a differentiated chemical in the brain - "chemical memory pellets", as it were' (1973:5-6).

In the same article he goes on to say that there is every reason to believe that storage of memory in the brain is a

> 'many-splendoured, multi-phased actively changing affair. That is, any single memory is not merely deposited in a completed form in the brain. Rather, it goes through a complex developmental history in the brain in which it changes from a ST to a LT (long term) memory. And each stage in this consolidation process, seems to be dependent upon different though interrelated chemical mechanisms. First we can assume that immediately after every experience, a relatively short-lived reverberatory process is set up within the brain. This process continues for a time after the stimulus

disappears and permits us to remember events which occurred moments or minutes ago. But this reverberatory process fairly quickly decays and disappears, and as it does, so does the related memory. However under certain conditions, the short-term reverberatory process, before it disappears completely from the scene, triggers off a second and quite different series of events in the brain. This second series of events involves the release of new RNA's or the production of new proteins and other macromolecules. And these chemical changes are relatively long-lasting and serve as the physical bases of our long-term memories' (Ibid:6).

Krech has argued that if the robustness or survival of the initial reverberatory process could be increased, there would be a greater probability of converting the STM to LTM. Experiments using various drugs on rats have yielded evidence that it is possible to affect LTM without affecting STM, thereby producing additional supportive evidence that the two kinds of memory are indeed different mechanisms and not part of a memory continuum as had been suggested in the past (Norman:1968).

The implications for reading, and indeed learning in general, are indeed fascinating, and suggest that forgetting could be attributable to failure to strengthen STM by keeping memory active by thinking or reasoning, or other activity, thereby increasing the chances of conversion to LTM.

One of the implications of limited processing

capacity as we saw it was that if the reader could develop a degree of automaticity at some level where this was possible, there would be greater degree of attention available where it was most needed, at the reasoning level. At lower decoding level where rules functioned, processing could be almost habitual, proceeding along well-run lines like a frequently-used routine or programme.

It also implied that the non-native reader by using his memory and processing capacity at the perceptual level, could not spare the attention or time necessary for obtaining meaning at the discourse level; and where input continued for many minutes, the memory was likely to become overloaded unless efficient processing occurred. If he was concerned with the problem of deciphering or decoding, he would have little opportunity to assign meanings to larger units so that the quality of items stored in LTM was likely to suffer.

We know from Simon (1969)[1] that transfer from STM to LTM seems to occur only once in four or five seconds which implies that the duration capability of STM is always used to the maximum. This with efficient reading should allow for re-organization and chunking of input to help reduce the amount to be stored to an efficient minimum. The time in STM also implies that it is possible for most verification or matching tasks that depend on fairly short

---

1. H.A. Simon quoted by Holmes and Smith (1973:64).

and self-contained units, to be done here without any need to enter LTM.

Those theories that suggested that immediate short-term and long-term memory storage devices paralleled levels of processing like letter identification, morpho-graphological, grammatical, syntactical and semantic, are obviously untenable, since we think now that all levels of processing can occur in immediate, short or long term memory. However, if processing of data is slower than rate of input, some of it will either decay and be forgotten even before it is processed, or it will only be partially processed, perhaps identified, but not comprehended. And in this partially understood form might well be stored in LTM. However, most adult readers, we would assume, would prolong fixation time or cut down input to allow what has been seen to be processed and understood as far as possible. From this discussion it is obvious we viewed organization of memory capacity as fairly crucial to the entire cognitive processing system, and comprising three types.

4.6.4 Components in Storage

Hitherto we have been looking at various parts of the system and their relationship to incoming data. But in order for any meaning to be assigned, it follows that knowledge of the foreign language along with knowledge of the first language, the world, past experiences etc, be stored somewhere. The obvious place is the long term memory. So following the Ruddell and Venezky/Calfee models

we suggest that all concepts and knowledge, and linguistic components must have a place here.

Among other things, the store would include knowledge of:

(i) features needed for letter identification in English.

(ii) features and rules governing morphographemes

(iii) features associated with form classes

(iv) lexical items of high and low frequency

(v) associational rules of the type found in case grammar where lexical items are related to each other in meaningful ways.

(vi) rules of grammar

(vii) rules of syntax and word order in English

(viii) composition of phrases of different kinds

(ix) clause and intra-sentential relations and patterns

(x) sentence patterns

(xi) inter-sentential relations

(xii) rules of discourse and features of identification

Because we have taken discourse as input, the above items will be represented in long term store in terms of form, function and meaning. In addition, where forms have names or labels associated with visual or acoustic representations, these too must be stored if learnt, since they are often used in learning situations.

These then are some of the details our model will

be taking into account. Because we are concerned with a second language, the implication is that the reader already possesses a vast storehouse of knowledge in his first language, and to a lesser extent in English. Knowledge of the world, including facts of the physical system, time, space, weight etc. are already available to him and taken for granted.

But we assume that what is different in a reading model designed for a second language is primarily the form and function, and not the meaning which is shared with the first language. Concepts and ideas are related to meaning and are not often language specific. This is why we adopted the definition of comprehension referred to in Section 2.3. The reader can go from English to meaning and back to another language making translation possible. Obviously the verbal units corresponding to the meaning may not occur in a one to one ratio, but as far as we know, most concepts if they are not shared, can be created through verbal ideas. It is also possible that 'some information which is handled semantically in one language is handled syntactically in another', Olson (1972:149).

But the common meaning component was demonstrated in Kolers' experiment with bilinguals (1970:111). Students of English and French were presented with connected discourse in the two languages, and the ease with which reading was accomplished and the frequent use of forms different from the printed text led Kolers to postulate that reading was only incidentally visual and subjects, who knew the language,

perceived the concepts and meanings rather than the forms. The extent to which meaning can be abstracted by the NNS will ipso facto depend on the degree of competence in English.

4.7 SUMMARY

Our brief survey of some of the more popular theories and models of reading has revealed the differing areas of concern that have attracted the attention of reading specialists over the past decade. And in the attempts to explain some of the component skills that might be involved when reading is in progress, we have come across a proliferation of terms, some distinct and some overlapping. We have had 'input' and 'output', 'sensory' and 'perceptual' data, memory storage devices with varying capacities for holding information, and components of the memory store, like lexical items, syntactic and transformational-generative rules etc. We have seen that some of the information included in the theories and models is based on observable phenomena which therefore is partially factual, while many of the views pertaining to the cognitive processes are necessarily hypothetical.

But the value, primarily of the four models reported, lies not so much in the differences as in their complementary nature. This is why, as we explained in the implications for our model, we will be using ideas from each of these four models, trying to do what Ruddell attempted, synthesise information into a coherent system but in our case with the benefit of his version and subsequent experimental findings to help us.

SECTION THREE

THE READING MODEL

CHAPTER FIVE

THE READING MODEL

5. GENERAL REMARKS

Since any proposed task that seeks to measure this elusive ghost called 'comprehension' must reflect to some extent the hypothetical construct on which it is based, we propose in this Chapter to present a reading model for a second language. It is of necessity speculative, but it is hoped that this attempt to integrate facts, results of experimental findings and ideas into a comprehensive but consistent framework is a major step in the right direction. We have utilised views and opinions already referred to in Chapter Four, and propose a theoretical model that incorporates what we feel must be accomodated in any explanation of the reading process, especially when the user is a non-native speaker of the language in which the information is presented.

We have taken as our starting point, printed matter in English read by a NNS, and the model traces the activity involved in the information abstracting process. We adopted the view suggested by Kingston (1960) that a model is a form of reasoning and presents a preliminary stage of thinking and will be open to alteration and expansion. Although our model cannot hope to solve or

explain all the many problems concerning reading, it has the virtue of trying to combine known information into a broad conceptual framework. It is a working approximation of the more formal theory. Since our aim was to investigate reading activity, we included those details that appeared to contribute more directly to this end.

5.1 THE READING MODEL

Our model differs from those reviewed in Chapter Four by being more detailed and comprehensive, including stages often ignored. It represents a synthesis of ideas and information gleaned from widely differing sources, and while recognizing that the multi-dimensional nature of the reading process is only vaguely realised, seeks to present a simplified version of an incredibly efficient mechanism that functions despite lapses and temporary breakdowns. Since our model sought to account for discourse comprehension at academic level, we used as input texts of 1000 words or more.

From Section 2.3 and the criticism of some of the literature reviewed in Chapters Three and Four, it is evident that for us comprehension includes activities that involve thinking, reasoning, organization and evaluation. Reading in our context was viewed as largely instrumental, and the function of the activity was primarily to permit the NNS to abstract the information in the text and interpret the contents for their meaning value. Success and failure

was largely a question of how well this purpose was achieved and since in our context, time was a valuable commodity, we were concerned with the speed of comprehension, the rate at which the NNS obtained understanding and how best we could measure this. We present below the details of our model which for convenience was divided into five stages.

Stage One : Input

Stage Two : Conversion. Here visual images of printed matter are converted to a code or form more convenient for cognitive processing.

Stage Three : Cognitive Processing for meaning. This stage represents the end of the information abstracting journey.

Stage Four : Re-conversion. Since in our model verbal evidence was required of the comprehension of the material used for input, this understanding has to be re-converted to a form more suitable for response generation.

Stage Five : Output or Response : This though determined by the tester and the task, involves the production or recognition of a response that is then matched against responses pre-judged as correct and acceptable in the context.

In our description we have avoided many of the pitfalls evident in the other models by endeavouring to incorporate answers to the following three questions at each stage in the process:

(i) What are the components involved?

(ii) What are the processes entailed?

(iii) What are the limitations they are operating under?

The answers, as will be seen, vary in depth of detail from one stage to the next. As we said earlier, the information reported is partly factual, partly inferential, and partly speculative. In some cases no answers have been attempted but it has still seemed useful to raise the questions if only to remind ourselves that failure to provide an answer doesn't warrant our ignoring their existance or their relevance to the model and the processing.

5.1.1 Stage One : Input

We begin with the NNS faced with English text. The eye when looking at print fixates for about 250 msec. during which what is within the field of vision can be imprinted on the retina; and this movement occurs on an average about four times a second. The retinal image stimulates neural activity necessary for the processing to begin. We also know that the field of vision covers an area, oval in shape and roughly two inches across and one inch down. Within this area the focussing occurs, so that part of the field can be fairly sharply defined; this latter area is only three letter spaces. The remainder though visible is less clear. The central area covers three rows of print about 15 to 20 letter spaces,[1] that is,

---

1. Philip Gough (1972:333). See also G.Sperling (1960) in Psychological Monograph 74 : The information available in brief visual presentations.

about two words in the central line with some print in the lines above and below this. What the retinal image obtains is a multi-dimensional image including letter or word-shapes, spaces, details of print-size, clarity and thickness of print, spatial distribution of print including punctuation, capitalization etc.

Within the eye there is a focussing mechanism which takes about 100 msec. to refine the image, and which permits certain features to be highlighted so that the imprint itself reflects selectivity. We assume that the eye focusses in response to a motor direction, and although in the first language this is probably developed to a level of automaticity not all NNS may have this ability. It is, however, amenable to control. The eye then has a sensory input system and a motor control mechanism.

If reading involved the use only of eye and focus mechanisms sans processing, reading speeds would be fairly stable for first language readers at around 2·5 x 4 words a second, or 10 x 60 = 600 words a minute. Second language readers would probably be slower, say at 300 words per minute. In actual fact we know that the latter speed is considered 'good' for first language users. Most of our readers will have speeds below this in English, with a few at this range. This implies that the reader deliberately controls input to permit processing to occur. Rate of input can be reduced either by longer or fewer fixations, smaller units of input, larger overlapping of visual field from one fixation to the next, or by a combination of all these

devices and strategies.

We take then for granted that there is more to reading than allowing the eye to proceed from left to right and down the page. One further proof is the common experience shared by all readers of occasionally progressing through a text and suddenly consciously realising that we have not 'taken in' a word of the text. We have been going through the motions of reading automatically, but because our attention was directed elsewhere, the input has not been processed for meaning, and has been lost and forgotten. So we consciously direct the eyes to regress and re-locate the place in the text where we left off attending and re-read again, re-inputting the data for processing. This awareness is the result of our ability to monitor our reading, and more will be said of this in Stage Three of the process.

All regressions are bound to slow down the reader, but they are a necessary safety-device to be used when needed. Since attention capacity is limited, we cannot always be paying attention and regressions can help. It follows that it will be used to advantage when we need to resolve doubt and uncertainty, correct mistakes or re-confirm earlier input; but a reader who has constant recourse to regression will interrupt his own reading to the point of interference.

We have then referred to the eye mechanism, the limitations it works under, and the processes involved. But we have not said much about the third component, the

input, which serves as the stimulus that triggers off the entire operation; nor have we said much about its size. That it can vary from one reader to another we have made clear, but we assume that even for any one reader the size of the input varies. It is not a constant. It will be the largest manageable unit - LMU (Venezky and Calfee : 1970) - that the reader feels he can cope with, and as reported by Schlesinger (1968), it is likely to coincide with a syntactic unit if the reader is reasonably competent. 'Unit' then refers to the linguistic or verbal elements that are undergoing processing at any one stage. At visual level, it may be a letter, a word, or a syntactic grouping, if this coincides with input. But once the unit has entered the system, the make-up and size will be the result of processing together with the cognitive strategies developed by the reader.

We wish here to make the point that the formation and size of the unit will change as the verbal data proceeds through the language processing system, becoming progressively more loaded with information. The unit in working memory will be the condensed core of perhaps a dozen units input at Stage One through the visual mode into immediate memory.

Operating under the limitations imposed by span of absolute judgement, the reader by habit, will either narrow down or expand his input to exclude or include data from peripheral vision. The skilled reader will increase distance between fixations, avoiding overlap, and will rely on his relative judgement in guessing what the expanded

input is likely to be, knowing that he is right most of the time. The NNS who is uncertain of his knowledge of English will on the other hand, wish to be absolutely certain that he has identified each item in the visual field, and will tend not just to refuse to process the peripheral data, but will direct the next fixation to a point that provides a large overlap of the area under focus. He will cut down distance between fixations and therefore limit input, reducing the speed of reading.

The proficient reader reads using information from what he sees and, as Goodman reminded us, what he expects to see. 'The peripheral retina plays an important role in the normal reading process by providing direction to successive fixations and fragments or partial cues; and knowledge of language structure enables the fragments to be productively used in an anticipatory way' Schiffman (1972:610). Input then is the result of what Hochberg and Brooks (1970:311) have called PSG - peripheral search guidance - acting under the control of CSG - cognitive search guidance - and the make-up and size of the linguistic unit will reflect the familiarity and ease the individual reader has in processing the text.

The input must be processed fairly rapidly since we know that if the visual source is cut off, we retain a retinal image or memory trace for a brief second or two before it fades. Since normal reading fixations occur at intervals of roughly 250 msec. the image must have been

transferred from the retina by the end of this period to leave the field clear for input of new visual data.

Our model then provides for variations in size of LMU's resulting from individual differences in competence and reading strategies. Speed of reading will therefore be one of the major differences between NNS who are familiar with English and so rely on relative judgement to increase rate of input, and those for whom the language is still a very foreign and unfamiliar source of information, who because of their insecurity will feel the need to be absolutely sure of what they see before proceeding further. To summarise then, the rate and amount of input, and the size of the unit processed, are constrained both by the physical limitations of the visual mechanisms in use, and the cognitive controls exercised by the reader to ensure that he has the time and capacity to interpret what he reads for meaning before proceeding further.

5.1.2 Stage Two : Conversion

This is a stage hardly ever referred to in the models because we have little or no idea what it comprises, nor indeed the limitations it may have to take into account. But we do know that some process of conversion occurs. Stage One makes it very clear that the visual image decays. Since in reading, the data comes at continuous intervals, and the activity may continue for more than thirty or forty minutes or more, the retinal images must be speedily transferred if input is to be preserved for processing. It

is suggested that these images are converted into a code more conducive for cognitive processing. Since verbal input can be either visual, acoustic or tactile - as for braille - it seems economical to postulate that all verbal data undergo the same cognitive processes, have access to the same memory stores and utilise the same devices and mechanisms to abstract meaning, regardless of the source and mode of input. The input coming from different sensory locations using different sensory apparatus, carried by different neural chains must all be capable of conversion to the same code in the interests of economy and efficiency.

It seems imperative to ensure that our model of reading does not exclude the possibility of use by other input modes. The implication therefore is that conversion to this common code occurs before cognitive processing begins, either <u>en route</u> to the language processing centre located in the cortical tissue, or on entry to it. Of the two alternatives, conversion on entry to the system seems the better choice. Our non-native user has learned to recognise what constitutes verbal data in English so that this is the part of the input that is forwarded to the centre. But this recognition has had to be learned; and the degree of automaticity exercised cuts down the decision-making time leading to faster processing.

But prior to this, when the script was still unknown - as with Greek, Arabic or Indian learners exposed to English - the visual forms would be treated as visual

imagery and sent to the visual processing centre. As soon as verbal matter can be differentiated, those parts of the input that are judged to relate to language are forwarded to the centre for conversion into the new code.

Perhaps this also explains why unskilled readers often fail to take non-verbal data on the page into account, and do not process it for linguistic meaning. They treat position, partitioning, spacing etc.,as visual matter rather than as meaningful signals used by the conventions of writing and printing in the foreign language, and having a linguistic function. Once again the implications are that the NNS needs to learn that these devices are useful because they convey meaning. The efficient reader by maximum utilisation of selective input, is lines ahead of the inefficient reader who regards everything except letter-shapes as visual interference to be ignored.

The visual - or auditory, or tactile - centre selects from the input unit the data that needs further processing in the language centre, and retains or allows to fade non-linguistic data. This decision is the result of learning by association.

To summarize then, in Stage Two the data selected by PSG or peripheral search guidance under direction from CSG or cognitive search guidance is now forwarded to the visual centre. The unit is further reduced and only linguistic+verbal data is forwarded to the language processing centre. On entry, it is converted to a common code shared by all modes of verbal input and the components in long term store. The input unit is now ready for

processing for meaning. With constant usage, it is very possible that with entry of verbal input, a program is triggered off whereby the unit by-passes the visual centre proceeding directly to the language centre, again cutting down on precious time before processing can begin. The better the reader, the more he understands what he 'sees', and the less detail does he need in order to identify forms which can be interpreted meaningfully.

5.1.3 Stage Three : Cognitive Processing System

This stage constitutes the heart of the language processing system and is described in varying degrees in most models of reading. But before beginning our description, we would like to put forward for consideration a view that does not appear in any of the literature. We would like to suggest that the entire language processing system is capable of directing organization differently when either perceptual or motor processes are involved.

In a lot of the current literature on linguistic theory there appear to be two views regarding the organization of meaning in terms of case grammar. We do not agree that one is correct; what we suggest is that both views as presented by Fillmore and Chafe are complementary, both having psychological bias. When perceptual processes are involved in verbal matters, we agree with Fillmore (1968) that the meaning tends to be organized round the referrent, the nominal. In experiments, subjects have reported remembering the 'topic' or 'theme' -

to use these in a loose general sense - long after they had forgotten what was said about it, Gomulicki (1956). In reading therefore, the main concern will be in trying to establish firstly what is being talked about, and all organization of information will be directed by this central concern, the 'agent - action - effect' to quote Gomulicki.

But when response generation is entailed, we suggest that there is a psychological bias towards organization of information round the verbal grouping, the activity or action that the individual wants to relate. We suggest that the organization of linguistic items in store are capable of change and that a motor response causes the system to adapt to a new direction, so that Chafe's (1970) approach with the verbal element dominating and dictating the choice of accompanying elements seems particularly apt for this purpose.

But to return to our model, we will be taking many of our ideas from the Venezky/Calfee, the Ruddell and the Goodman models already reviewed. This stage is concerned with the interpretation of the already coded input for meaning. It is here that the limitations affecting attention and processing capacity operate.

5.1.3.1 Components involved

The components which are involved in the processing and which will be described in greater detail later are:

(i) MEMORY - with three storage facilities differing both in capacity and durability.

(ii) ITEMS - contained in the temporary and permanent stores comprising linguistic, experiential and world knowledge etc (See Section 4.6.4).

(iii) RETRIEVAL MECHANISMS - permitting access to, and use of the items in permanent and temporary storage.

(iv) CONTROL UNIT - this enables the reader to monitor his own reading activity by:

(a) distribution of attention capacity according to the necessity for decision-making at different times during the processing activity.

(b) distribution of processing capacity thereby activating the various operations that need to be performed at the various levels including response generation for output.

(v) FEEDBACK MECHANISMS - operated by the control unit which enable it to:

(a) perform the functions listed under (iv).

(b) have access to and from different parts of the system.

(c) monitor input by directing the focussing mechanism at the visual level in Stage One of the processing (Section 5.1.1.).

(vi) COGNITIVE STRATEGIES AND SKILLS - developed for use in reading and interpreting verbal matter either through training or experience. This is viewed partly as similar to programmes used in computer technology by which the reader can set in motion certain procedures that are carried out with minimal attention and direction e.g. a perceptual strategy versus a response strategy, detailed reading versus skimming etc. In addition, the use of strategies and skills enables the components in stores to be re-organized to create new forms of meaning. It is this ability that makes for the generation of original and creative ideas under stimulus from printed input.

5.1.3.2 The Process

The processes that occur are in general three-fold:

(i) Recognition - by which we mean that the input is recognized as a linguistic component and a unit is then created for processing. This unit will be the LMU the reader can create aided by familiarity and exposure to the language of input, that is, English.

(ii) Identification - by this we mean that the language component that has been recognised and formed into a unit, must now be matched against already existing items in memory stores. Retrieval mechanisms are activated and a search begun for matching. Once the LMU has been tagged as equivalent to the item in storage, the next part of the process begins. It is necessary that form and function be identified.

(iii) Meaning assigned - This stage incorporates the assignation of the meaning associated with the unit in permanent store to the LMU identified.

As we said in our discussion of the linguistic items in permanent store, we have information regarding the formal properties of items which helps with identification. We have meanings associated with these forms, and this is the meaning that is generally learnt and assigned by the unskilled reader.

But it is also possible that knowledge of how these forms function has been stored. And these functions will also have meanings that can be assigned. This is the meaning the competent reader will assign.

The meaning then will vary according to whether the form alone or the function has been identified and what richness of meaning and concept is associated with this

item in permanent store. We appreciate that a lexical item like 'civilisation' may be identified with concepts as varied as art, scientific development, painting, Western countries or Eastern thought. What the retrieval of meaning will do is set up or activate an 'area' of meaning and the extent to which incoming text coincides with this area, processing will be facilitated because predictions will be fulfilled. It is however, just as likely that some of the area will be eliminated by reference to later parts of the discourse. 'Meaning' then like the 'unit' is not a static concept, but a dynamic one where the interpretation assigned in one part can be altered and refined as the discourse proceeds. The longer the text, the more likely that alternatives will be gradually reduced to the one meaning best fitted taking into account all the increasing constraints imposed by context. These then are the processes necessary for conversion of input into meaning, and which are being simultaneously performed on a succession of units in a continuous, cyclical process.

5.1.3.3 Definition of some terms

One of the main functions of the system along with meaning-abstraction is the reduction of the quantum of data by shedding, selectivity, and condensation, so that what is eventually processed for storage is the minimum necessary for the purpose for which the reading was undertaken.

This is essential for efficiency. One of the ways by which reduction can occur is by recognition and identification of aspects, some representing properties, others the items themselves, bearing the same meaning and which therefore are judged unnecessary or redundant. Redundancy has been defined by Whatmough (1958:77) as 'the presence of more clues than are strictly necessary'.

As the unit grows in size and complexity, processing activities will obviously refer to ever-changing data. To help us describe the cognitive processes we will be using the terms 'higher level' and 'lower level' processes. In much of the literature on reading theory, terms like 'higher level processes' and 'lower level processes' have been freely used without definition, sometimes referring to levels of linguistic analysis and on occasion to cognitive processes. The distinction has often been blurred almost as if linguistic levels were psychological concepts and the same thing.

Some of the other terminology associated with the ideas has come from theories of grammar and words like 'surface' and 'deep' structure once again are used by writers who fail to clarify how they are to be interpreted. Case grammarians have tended to use 'deep structure' to refer to meaning, using the 'surface structure' as representing the realisation of this meaning in language forms constrained by rules and conventions of usage. Transformational-generative disciples have used the terms

as two levels of syntactic form, separate from meaning e.g. Chomsky (1965:16) 'the syntactic component of a grammar must specify, for each sentence, a deep structure that determines its semantic interpretation and a surface structure that determines its phonetic interpretation.'

For the purposes of our model, we have separated form, function and meaning in the components held in permanent memory store. We recognise that with form there is a continuum representing in hierarchical ordering, linguistic elements ranging from letter shapes at the most elementary end of the cline to a grammar of perhaps rhetoric or discourse at the other. But for purposes of convenience we have divided the forms along with function and meaning into two levels : lower and higher. We use the term 'surface' or 'lower' level to refer to forms that reflect linguistic units from letter shapes to words and word-groups, involving their grammatical relationships of the kind exemplified by : article + noun, adjective + noun, phrases like 'as ... as', etc. This division into two levels reflects the view that at the lower end, the components are not only more strictly rule-governed, but also tend to occur in highly re-occurring patterns. They are therefore far more predictable, and so more easily learnt. And by frequent application of rules, a fairly high degree of automaticity can be built up in associating form, function and meaning. Thus at this lower level, processing can be faster because of

already established habitual links so that with the limited durability of input in immediate memory, it is processing at this level that occurs. 'Lower level processes' then is taken to refer to processing of lower level linguistic components and will be performed, whenever possible, in immediate or short term memory using minimal attention and processing capacity. Units here tend to conform to expected forms of usage, and being predictable, are processed with ease.

With larger linguistic components such as phrases, clauses, sentences, paragraphs etc, there is an increasing complexity of relationships possible between different parts of the verbal data. These larger syntactic structures have syntactic relationships, and when these are recognised and identified, the semantic structure of the same unit must now be mapped on to whatever represents the meaning. According to Simmons (1972) the syntactically related constituents must now be transformed into logically related meanings. An attractive explanation of how these relations might be meaningfully related is offered by case grammar.

'Each verb and each noun then contains within it, most importantly and above all, a lexical unit, which carries the main burden of the information conveyed by the verb or noun ... Each lexical unit has associated with it certain inherent features ... and certain contextual features ... Given a particular lexical unit, its inherent features are predictable; they are inherently associated

with it. We might say that they correspond to pieces of knowledge necessarily associated with the knowledge conveyed by the lexical unit. The same is not true of the contextual features, which either are chosen freely by the speaker to communicate whatever he wishes to communicate or, probably more often, are determined by contextual or discourse factors' Chafe (1972:43). The effect of these contextual rules is seen by Chafe in the addition or deletion of certain kinds of transitory constraints which will come and go as the discourse proceeds.

The relations established between the grammatical and semantic units cause the correct interpretation of the higher level components to be made depending on the richness of the network of links between items in the store possessed by the reader, and the efficiency of retrieval mechanisms.

At this deeper or higher level, the possibility of rule-governed patterning decreases and with it the degree of automaticity possible. Since this level reflects the individuality of the writer, more time, attention and processing capacity will be needed for recognition, identification and ascription of meaning. 'Higher' or 'deeper level processes' then refer in our model to the processing of higher level linguistic components, and these are likely to occur, if possible, in short term memory, but more often in working memory (to be explained later) which has greater temporal durability.

Just as the division of linguistic components into higher and lower levels has been made for convenience of description, so too, with the higher and lower level processes. We recognise that they too form part of a continuum with some performed faster because they can become habitual, and with other processes needing more time to organize. Units must be recognized before they can be identified and interpreted in a meaningful way taking contextual constraints into account.

The reader who has learnt by experience and training to reduce claims made by lower level components, thereby preserving attention and processing capacity for analysis at the deeper level, can increase speed and efficiency because he will be concentrating on meaning rather than decoding. The latter term is used as it normally is in studies and research on reading, as representing the recognition, identification and naming of forms and their pronunciation.

Since reading aloud will not be needed for our purposes, we have not included labelling of forms as one of the major components, but it will obviously exist and be used if reading comprehension is to be judged by the reader's ability to produce as output a vocal duplication of printed input.

Let us trace input as it quantitatively becomes lesser while increasing in density of information, as it proceeds through the cognitive processing system <u>via</u> the memory stores.

5.1.3.4 Memory

We would like to suggest that memory has three sub-components, each having storage facilities of varying capacity and duration, each imposing its own limitations on the input and processing that occurs.

5.1.3.4.1 Immediate Memory

IM or immediate memory receives the input as it enters the cognitive system. This has limited storage capacity and items can only be retained for a duration of one second. The amount of input may be seven, plus or minus two, discrete items if they are unconnected; but since our input is part of continuous discourse, and our reader not totally unskilled, the amount of input will really be controlled by rate of input and amount of intake during eye-fixation. We are told that the average reader needs four to six fixations to read a line of print. The input from four of these, coming one every 250 msec., can be retained in IM for processing, so that recognition of short grammatical units can help 'chunk' them together for identification and assignment of meaning, even if this is only partial. Much then depends on the size of the input. Retrieval mechanisms are now activated and engage in a search process for the purpose of matching the unit with data in store. If more information is needed, the half-processed or unprocessed unit is temporarily stored while more input is awaited. The addition of more items,

often bilateral context, may help in recognition. Once again permanent stores are searched by retrieval mechanisms for a match so that identification can occur. As the unit increases in size, more than one kind of processing will occur simultaneously resulting in a syntactical appreciation of the unit, plus an awareness of the grammatical relationships between the parts, and meaning of the lexical units if known. This implies that simultaneous retrieval mechanisms are activated.

Figures 7 and 8 present a schematic representation of the content and processing in immediate memory and may help to clarify the working of the system. Because of the time limit, only short units can be processed here but short sentences could easily be interpreted for meaning. (See Page 128).

Especially with foreign language users, different aspects of the store may well vary since the language system is likely to be unevenly learned. The LMU will vary from person to person and from time to time. Specific items within a LMU may be unknown so that the meaning assigned will reflect the individual's competence in dealing with the foreign language. An individual may possess a large vocabulary and still be unable to assign meaning to a grammatical or syntactic unit, while another may recognise the meaning of the LMU while not knowing a particular word in the unit.

Along with Venezky and Calfee, we would like to suggest that words of high frequency usage which have been input, stored and retrieved on many occasions, are more easily accessible than others. Use has established a multi-dimensional network linking them so that when an already existing network is used, processing time is saved. It is likely that words of low frequency use are difficult to retrieve because they have fewer interconnecting networks. And unless the retrieval stimuli is similar to the one that originally occurred in input for storage, the chances are that there will be a delay while the search in permanent memory seeks to trace the associations to obtain a match for the item.

But for our NNS, what is frequently used, and therefore, easily accessible, is more probably determined by classroom conditions than by the dictates of the language per se, as is more likely with a first language. Items commonly occurring in the language, even if recognised, may not be correctly identified in the context since many lexical items are polysemous, and the reader may well hesitate to assign meaning. What is hence frequently used in the language is not necessarily frequently used by the NNS who has up till now had restricted exposure to English, and lower level elements may continue to offer a problem despite years of study.

We now have input, chunked or not chunked, processed or partially processed proceeding to the next store since it will be displaced after one second by incoming stimuli. Because IM is needed for more data, the contents are forwarded to STM or short term memory, and we suggest, some information regarding the processing level is sent back to the control unit.

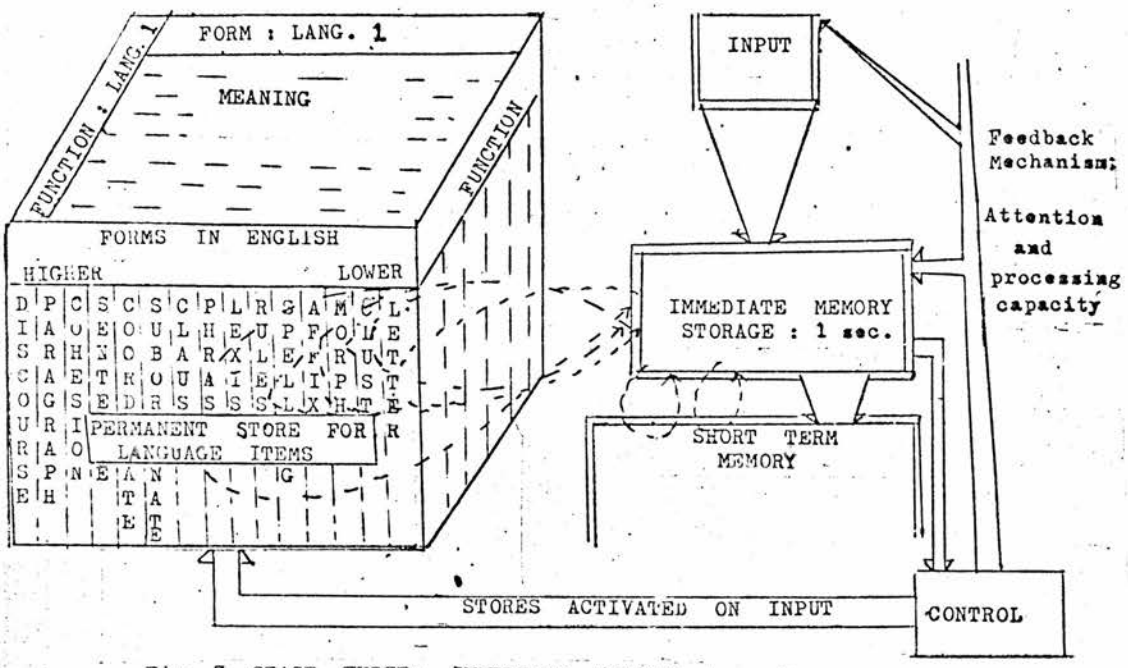


Fig. 7 STAGE THREE: IMMEDIATE MEMORY

Retrieval mechanisms which help identify form and function of the unit, so that meaning can be assigned.

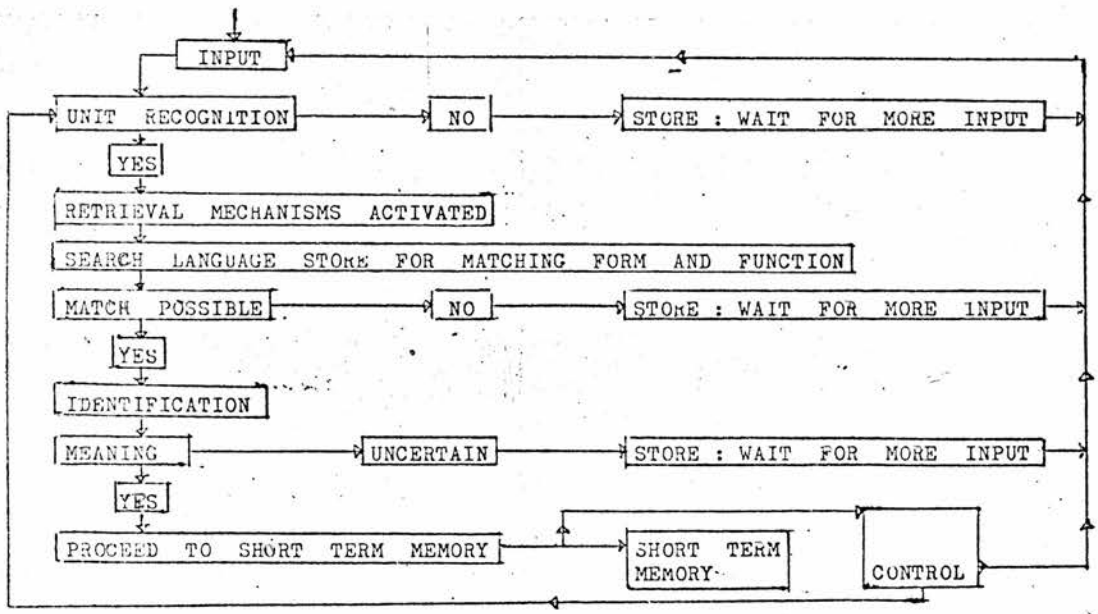


Fig. 8 STAGE THREE : SCHEMATIC REPRESENTATION OF IMMEDIATE MEMORY

5.1.3.4.2 Short Term Memory

STM is probably part of the IM continuum, where the storage capacity is still limited by the number of items that can be retained. Although there is no restriction on size of item, there is a maximum durability in time of four to five seconds. The five to seven units that can be stored for re-chunking and processing may represent the input of one or more sentences, if short. It is here that larger units of syntax can be processed and most intra-sentential links recognised, identified and interpreted by the average reader. Higher level processing which may already have begun in IM, is continued on in STM and if input is limited may be completed, so that meaning is available for use in response, recall or other activities. The simultaneous processing of units of varying size and complexity, the simultaneous search and retrieval of items from permanent store and the simultaneous evaluation and meaningful interpretation while the input is in progress gives us a vague idea of the incredible complexity of the system.

As we saw, both IM and STM have temporary storage facilities where current input can be retained usually to permit chunking to occur so that new processes can operate. They both have access via retrieval mechanisms to the permanent store without which little or no learning and use of language could occur. We

suggest that IM and STM be looked on as dual-function sub-components having storage facilities and a working, processing area, both functions confined to current input in present time. The results of STM processing are, at the end of four to five seconds, ready to be passed on elsewhere since new data from IM needs to be accommodated. This data, considerably reduced - since all redundancy as far as is recognised is retained in STM and so lost due to displacement by incoming data - now enters working memory, WM. This we regard as the working area of long term memory, LTM.

5.1.3.4.3 Long Term Memory

Very little is said about WM in the literature although where referred to it is regarded as synonymous with STM, Mackworth (1972:728). But taking our cue from the Venezky/Calfee model again, we consider that one of the functions of WM is similar to their TKS, Temporary Knowledge Store. And one of the limitations is perhaps what Goodman meant but did not specify when he used Medium Term Memory in his flow chart of 1970 (Page 83). The WM is where all data about the text being read is collated, modified, re-organised and again reduced by elimination of redundant information. It is the contents of WM that will later be available for storage in LTM, and for immediate use in some task for which the reading has been undertaken, perhaps a summary, answering questions or as preparation for some assignment. It is WM which,

via the control and feedback mechanism, helps to guide and predict input at visual level. Data that is input must survive for more than the limited durability of STM, and reading may be an indefinite activity lasting for indefinite time. If the total input is limited, the temptation for most students is to store as much of the detailed information as the reader thinks he will be able to recall or may need for later use. But if the passage is long as most academic assignments are, the NNS must, if he is efficient, be far more selective in what he wishes to remember. He needs to build up a reduced, condensed version of the information content. This, as we said earlier, is one of the main functions of the cognitive activity in stage three.

Unlike IM and STM, we know that LTM has unlimited capacity for storage and for an unlimited time. There is however one limitation. Items cannot be poured into LTM in a continuous stream. The safety mechanism that operates, limits entry into LTM to only once in four or five seconds. This means that whereas storage in IM is optional, available for use if needed, the storage facility in STM is essential.

We suggest that LTM, like IM and STM, has a dual function : a storage capacity and a working area, and our WM is the working area of LTM. The difference between the two functions however is that what is stored belongs and is the result of past time. To activate

any part of LTM, retrieval mechanisms must be activated and relevant items brought to WM which is concerned with current processing. WM is activated in present time as are IM and STM, and so WM must share with IM and STM the limitations imposed by attention and processing capacity.

The degree of attention that can be given is controlled by the central monitoring unit, but we must not forget that if processes involved at the lower level have not been learnt to the point of automaticity, that is where a greater part of the overall attention will go. Since reading is a visual activity, what is physically input and how it is to be initially recognised will have first claim on available capacity. 'Thus if a reader requires considerable capacity to decode a single word, his processing capacity is less available for higher order integrated processes, for example memory for the just previously coded word may suffer, memory for the preceeding phrase may decrease, and the subject's ability to predict what he is yet to encounter on the printed page may diminish', Perfetti and Hogaboam (1975:461)

Working memory then has the advantages of capacity and durability associated with LTM, but it is temporary in that it depends for its existence on processing activities. As long as rehearsal strategies of one kind or another will keep the contents active, they will survive. This means that WM is created by

the central cognitive system for its own purposes to accommodate input either from outside stimuli or as a result of self-generative sources. It has properties of both temporary and permanent memory, permitting greater storage and duration for current thinking. The latter is what helps to keep alive the reverberatory processes that, as Krech (1973) suggested, are contributory to conversion of STM to LTM.

The NNS who uses his limited capacities in WM is the reader who has not needed to divert them to the lower level in IM and STM. And because they are not needed there, he can process the total contents for efficient storage later in LTM from where it can be retrieved when necessary.

The unskilled NNS who has no capacity to spare for higher level processes has less likelihood of activating WM and so of remembering meaning of larger units of the discourse. When fairly short units conform to frequently used patterns, processing will be quick; but when demands are imposed on the system by length, complexity, non-conformity or deviancy of some kind, then the time needed to comprehend will be greater, and those memory-level components that have limited capacity will become overloaded, and the uncomprehended language units may either be forgotten or committed to WM. If further processing occurs, the input may be understood, if not it may still be accessible for recall if stored in LTM,

but in the same state of incomprehension. Experienced readers often indulge in this practice, putting aside in LTM, matters for later re-consideration. This is only feasible if the language has been processed and the concept or idea conveyed needs re-thinking and understanding. Lexical items too may be stored for later use with a dictionary; but if most of what is read is not comprehended, what is stored will be so fragmentary and haphazard that it will have little to yield by way of meaning.

It would be inefficient and we think unacceptable in any theoretical model to postulate that all data in STM had to go into LTM for permanent storage. This is why we see WM as a crucial part of the memory system and disagree with both Geyer (1966) and Gough (1972) in ascribing to it the function of a 'buffer storage' a place where Gough suggested sentences go to when they are understood - a PWSGWTAU - existing between cognition and response. To have this type of storage would imply that total meaning was simply an addition of the meanings of individual sentences making up the discourse. This would permit most of the information redundancy to be stored.

We feel that the NNS who is unskilled and inefficient may be doing just this, using WM more for storage than for re-organization and processing because he is unable to be selective and so stores literal meaning

of units. We suggest that if WM is used only as a storage module, input here will fade exactly as it does in IM and STM because no processes are helping to keep it alive.

But 'meaning of an utterance is not a linear sum of the meanings of words that comprise it' Miller (1973b:16). Since even at the intra-sentential level, 'the syntactic structure of a sentence imposes groupings that govern the interactions between the meanings of the words in that sentence' (Ibid), we recognise that larger units of text are equally likely to exert influence on the way in which the individual parts are to be interpreted. At the grammatic and syntactic level, meaning is also abstracted from the kind of word order operating in the sentence in English, and the kind of ordering operating among the concepts in a paragraph, and the organisation of these into discourse. The multi-level layering of units are each meaningful in different ways and both forms and functions need to be recognised and understood for comprehension.

Sensitivity to these levels of meaning can only come when the sentence is judged in relation to the whole, and the NNS needs knowledge and skill in abstracting all the relevant cues that will help him evaluate what is important for retention and storage. He needs syntactic and semantic cues since both are necessary for language comprehension.

5.1.3.5 Skills and Strategies

The particular function that they play is in utilising the contents of stores and input to provide the solution to the task for which the reading was undertaken. They form part of long term memory and it is the quality of these skills that make for efficiency. All native speakers at university level are by definition proficient readers, but only some are efficient. We also feel that in most of the literature on models of comprehension, the latter is seen basically as a matching operation very similar to the processes outlined in Section 5.1.3.2. But if this were all, comprehension would be activation of static concepts in the sense that only old, already existing meanings would be available for reference. To enable the creative aspect of the cognitive system to function, we view strategies as an essential part of processing activity giving to understanding a dynamic and individual function. We view strategies as techniques or devices which if used frequently become skills requiring a lower level of attention and processing capacity to go into operation.

It is the strategies that give to long term memory its generative capacity so necessary for creative thought. It is this facility that enables a reader to draw inferences and evaluate text. The utilisation of knowledge stored by interacting with incoming text helps to produce new insights into the subject matter. This is

possible when the reader goes beyond the explicit language forms to make connections that the less proficient reader fails to see.

We do not dispute that intelligence is probably very closely associated with the skills and strategies developed by the reader. This is why we have maintained that our NNS, who has the intelligence necessary for academic work, can compensate to some extent for language deficit by use of these strategies. We also view this component as dynamic, and suggest that help with evolving suitable strategies in dealing with difficult texts in the second language, will pay off in more efficient use of limited capacities and language knowledge.

5.1.3.6 The Control Unit

This monitoring unit is what activates the system even before physical input begins. It provides the necessary 'set'. Conscious decisions must be made in evaluation of the data at all levels. The way in which the reader approaches the task, the strategies he utilises, are all basically linked via the feedback mechanisms to the control unit. Habitual ways of thinking will be activated and skill in reading will reflect the ease with which the cognitive system processes the text. Different readers will approach the task with varying degrees of flexibility. Some will habitually infer meanings of words and phrases from context if

necessary; others may hesitate to input further wasting precious time regressing to see if the meaning has been given, rarely proceeding further to see if it comes later or is inferable. The efficiency with which the information abstracting process occurs is monitored by the central unit which controls the distribution and division of attention and processing capacity.

We have maintained that feedback regarding the state of the processing is sent back to control so that it can direct the focussing mechanism to either regress, slow down or move faster. If response to the input is needed, control must decide how much share of overall capacity is to go to output, and whether a balance can be maintained. The presence of suitable strategies will enable the Control Unit to function more efficiently so that the two are inextricably linked.

5.1.3.7 Summary

As discourse proceeds and is input through IM, STM and to LTM via WM, the reader attempts to recognize, identify and meaningfully interpret the units which increase in density. The reader uses all his knowledge of English to help him process as efficiently and quickly as possible. The degree of automaticity existing for lower level processes will determine the extent of higher level processing that can be performed on the growing body of data in WM. And this will represent the minimum necessary for efficient remembering. In order to achieve

this efficient minimum, the reader has employed the skills and strategies he has already developed for use in cognitive tasks. His ability to utilise these in English will be essential for success. The control exercised by the monitoring unit enables distribution of attention and processing capacity - which are limited - to different parts of the system as needed. Feedback mechanisms that keep the control unit in touch with all parts of the system, also enable the reader to increase or decrease rate of input to allow the system time to catch up on its own processing. The ability to successfully predict reflects the accuracy with which reading has been comprehended and the ease with which this has been achieved. This prediction is aided by information from IM, STM and WM covering grammatic, syntactic and semantic information that imposes constraints on what can occur in the incoming text. The expectations built up are a result of knowledge both of the language in general and the specific text in particular.

The accompanying diagram in Figure 9 presents in simplified fashion the working of the cognitive processing system as we have attempted to describe it, and exemplifies what we defined comprehension to be. To requote: it 'involves the recognition of written or printed symbols which serve as stimuli for recall of meanings built up through the reader's past experience. New meanings are derived through manipulation of concepts

already in his possession.... The reading process involves both the acquisition of the meanings intended by the writer, and the reader's own contributions in the form of interpretation, evaluation, and reflection of these meanings', Bond and Tinker (1957:19).

By the end of stage three the reader has abstracted from the input his understanding of the text. But for our model, we need an output from the reader that will provide measurable evidence of his understanding.

5.1.3.8. Organization of Meaning for Response Generation

This is likely to be directly related to the task set and the purposes of reading. Questions based on, and asking for information connected with the text will rely heavily on the contents of working memory since no reader at this level will peruse the entire text for each answer he makes. He will be using what he has already abstracted and using the text only for confirmation or to resolve uncertainty, aided by his memory for location of place on the page for the necessary information that many readers retain (Zechmeister and McKillip:1972).

It is often at this point that the NNS finds it difficult to retrieve the necessary language forms needed to convey the meaning he wishes to communicate. The language is presumed available in the first language, but the degree of competence in the foreign language, English, will now be evident by the facility and correctness with

which meaning can now be re-encoded in the appropriate forms. If no form is easily retrieved, the reader must either omit response or present one in the only alternative available - his first language. The strategy that the reader will adopt, is now going to gear his knowledge towards a motor response. He will be attempting to organize output round the predicate element since the referrent will have been provided by the question asked. This implies that comprehension of the material read is bound to be at the mercy of re-encoding operations in any task that requires written evidence. If one wanted a more direct or 'purer' form of response one could present alternative forms in English from which the reader could select the best one he thought fitted the task. This latter task will involve input of the alternative forms to match them against data in working memory especially if the meanings were not explicitly stated in the text.

But written evidence involves two kinds of competence, both of which are relevant. One is competence in abstracting information of the material read, and the other, a competence in productive expression of the kind needed for the task.

We come then to Stage Four. The end of Stage Three marks the completion of the information-abstracting process as the diagram on the next page illustrates.

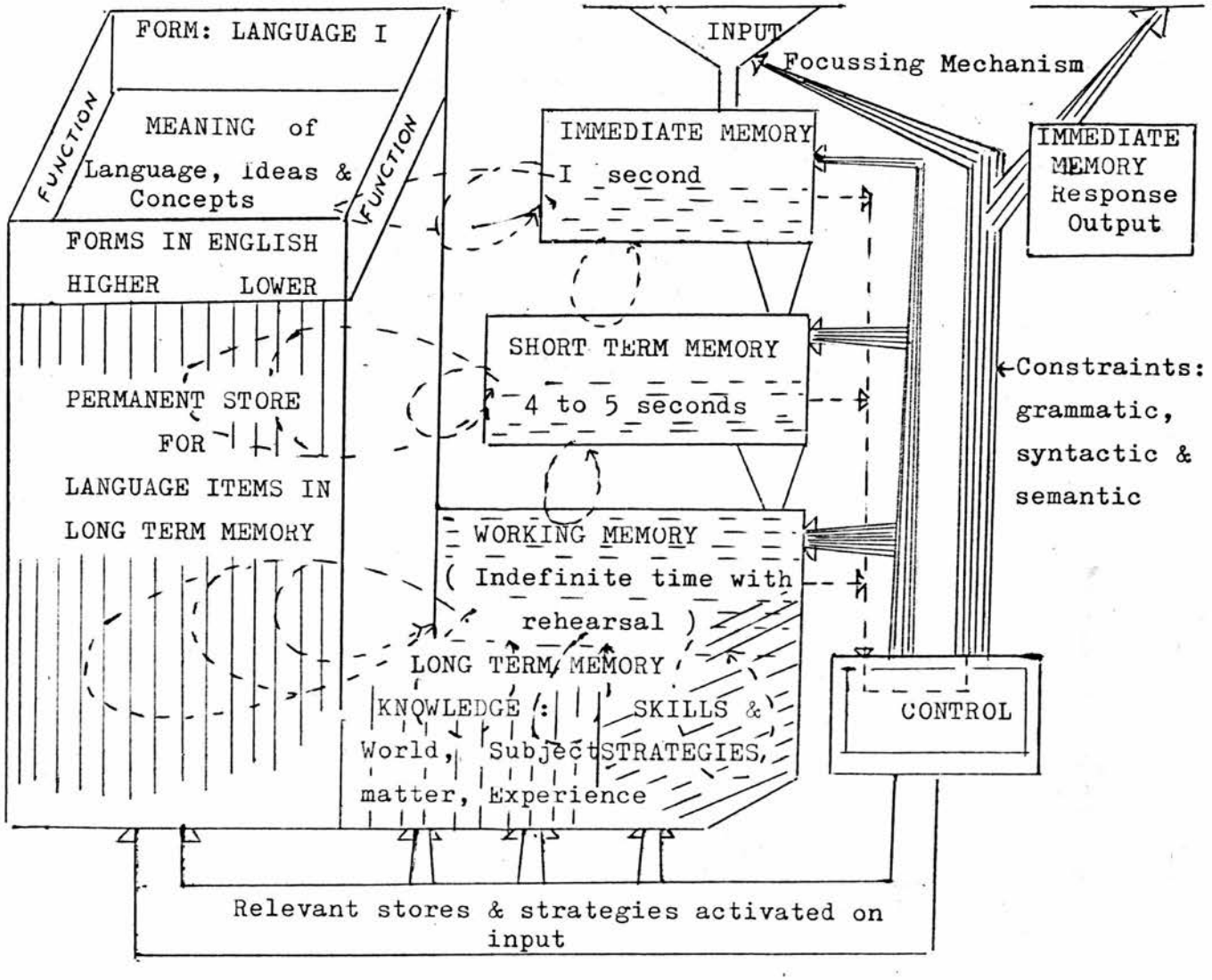
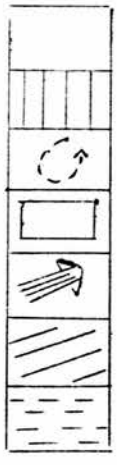


Fig. 9 STAGE THREE: MODEL OF COGNITIVE PROCESSING SYSTEM

Key :



Memory stores: Immediate, Short & Long Term Memory

Items in store

Retrieval Mechanisms

Control Unit monitoring entire system

Feedback Mechanisms to enable Control to function

Skills and Cognitive Strategies

Processing, organization, evaluation etc. of input

5.1.4 Stage Four : Re-conversion

In response then to the 'task-induced cognitive operation' to use Frederiksen's (1972) phrase, what the reader selects as the appropriate response may be elements, ideas, concepts from working memory aided by a search in permanent store for the associated verbal forms, or may simply be a matching requiring the concept 'same', 'different', 'right' or 'wrong' etc, to be output.

The answer activates a search in reverse since the forms must now be re-organised into grammatical and syntactic units that as far as the reader thinks best convey the semantic notion. Control monitoring will now have to divide capacity between two different operations, one which can best be described as an analytic break-up and the other which is a constructive synthesis. If part of the linguistic framework is provided, the search will also be constrained by the necessity of ensuring that the missing part fits the contextual constraints, i.e. the selected item will be 'tested' for goodness of fit, and the 'testing' will reflect the competence of the reader.

The selected unit which the reader considers best fits the requirements will be forwarded to the re-conversion unit. We would like to suggest that the response generated is passed into IM and on its way out of the cognitive processing system is re-converted to

a code or stimulus more appropriate for the kind of output that is needed.

At the re-conversion centre, we suggest that it is possible for either the visual or acoustic or sign mode to be associated with the verbal form, which is why we have made provision in permanent store for these specific forms to be stored. Since our model is concerned with the interpretation of comprehension from and into the written, visual form, this specific store will be searched and information relevant to the verbal form will be sent for re-conversion. What leaves the central cognitive system will then be a verbal response in the specific mode suited for output in a visual form.

5.1.5 Stage Five : Output

In addition to hand mechanisms that have been activated via the language processing centre, it is very likely that articulatory mechanisms too have been stimulated leading to sub-vocalisation or mouthing the response accompanying the writing process, particularly in the less proficient reader.

The reader has to balance a shift in attention and processing capacity, and may have to cut down intake to permit output to have its share of attention. Once again, the ease of retrieval of the correct form and the ease of production in physical terms of the form will reflect the familiarity the reader has both

with the language and the task. It will also reveal the organization and extent of his language store.

The speed and efficiency of the response must then be judged by the experimenter or teacher as acceptable or inadequate. To the extent that it is 'correct', the reader will be assumed to have understood the text, and produced evidence of his ability; to the extent that he cannot produce any evidence at all or has produced a response that violates some or all of the constraints imposed by the task, it will be inferred that he is less proficient. We suggest that tasks like summary-writing, comprehension questions or cloze are suitable.

5.2 IMPLICATIONS OF THE MODEL

Our model then has been reader-oriented and failure or success is measured in terms of how the reader performs the task for which the reading is undertaken, and the number of items successfully attempted will be taken as evidence of ability. In general we take the view that 'mistakes result from the psychological limitations of people who use the language, not from the language itself' Miller (1973:31).

However some constructions by their very nature impose severe demands on the temporary storage capacity of any component that attempts to deal with them. We have evidence that constructions of the 'onion' type

with self-embeddings are more difficult than left-recursive type, which are in turn more difficult than the right-recursive type. A simple example taken from Miller (Ibid:31) will make this clear.

The onion-type : The rapidity that the motion that the wing that the humming-bird has has has is remarkable.

Left-recursive : The humming-bird's wing's motion's rapidity is remarkable.

Right recursive : Remarkable is the rapidity of the motion of the wing of the humming-bird.

As we saw in our model, partially processed units can be forwarded to WM, but the delay is likely to interrupt input flow since attention will be focussed on difficult constructions. Knowledge of the language by the reader will encourage him to read ahead before he begins processing difficult structures, using bi-lateral context aided by visual and mental regression where necessary to abstract meaning. Constructions of the type quoted above are difficult because of the attempt to establish the identity of the referrent.

In continuous discourse, this same problem occurs in a different way. When the total input is short, contained in a sentence or paragraph unit, the subject of this discourse is likely to be identified with ease since it serves as a central focus for all the information in the text. Self-contained paragraph size units are generally on one theme. This is the kind of

MVD often used in testing situations as well as in classrooms. In longer texts where we have a number of paragraphs, the difficulty will lie in recognising and establishing the identity of the subject at various points in the discourse. A subject does not have unique identity that extends to sentence, paragraph and discourse. The development of ideas related to and following on from the general topic or theme may create problems of linkage where the reader will be expected to infer the connection. Because in English, repetition of the same form is often avoided for stylistic reasons, and other devices or lexical substitutes are used, the reader needs to recognise the referrent so that new information can be correctly organized round the appropriate claimant. Failure to do this will result in lack of coherence in the meaning comprehended. Moreover no writer presents all the various details that may logically be entailed when he makes two consecutive statements or moves from one idea to the next. He takes for granted that the reader possesses a knowledge of the world similar to his except in the local subject area that his material is concerned with. But even in the latter field, a certain knowledge is always subsumed under basic or general knowledge, and as the text proceeds, the more specific will be the expectations by the writer that the reader shares his knowledge.

Since presuppositions will become more passage specific as the discourse proceeds, the expectations that

the writer presumes will be shared by the reader will grow. If the reader has followed the thread of argument or development of informational content, he will be more able to predict both the language and the meaning as the text comes near conclusion. If he has not, the increasing length with its continued input in the foreign language is likely to make him increasingly less able to process the information, so that the gap between writer and reader is likely to widen.

Too often, NNS tend to compartmentalize information obtained from different sources. The concentration on the language of input has their exclusive attention and other stores containing related knowledge learnt through their first language are not activated. They lack the successful skills and strategies necessary to compensate for language deficit. The reader needs to utilise all the resources at his command to enable him to abstract the clues necessary for efficient storage of meaning. This is why we have maintained that the results in WM are the combined effects of language proficiency and the skills and strategies already possessed by the reader before he comes to the task. Obviously, the specific text will also help to determine ease or difficulty in reading for comprehension, but efficiency is essentially seen as a product of ability.

Our interest has centred on the process of information acquisition. How the input is processed and what the reader understands of the discourse while it

is and as it is being read, will provide us with information regarding individual ability, and evaluation of the contents of working memory will enable us to judge the efficiency with which the reader has reduced the input for further use. For the latter, techniques like questions, or summary writing, or application of the information to a problem exist, and obviously the extent to which they can elicit the contents of WM, to that extent they will be valid measures. But to attempt to infer information regarding the process is evidently a more difficult proposition. A technique that simulates, or more accurately, activates the process on a smaller scale will yield information relevant to the reading ability if we can find some method of measurement to evaluate the response.

It is suggested that the cloze technique is peculiarly suited to this task.

5.2.1 Cloze

We have tried to show that the system selects bits at visual level so that input does not overload the processing capacity, and one of the major functions of the cognitive centre has been the reduction of input at each stage. The central monitoring unit is able to direct the focussing mechanism under constraints imposed by information from IM, STM and WM, loosely coinciding with grammatic, syntactic and long range semantic context.

Failure at visual level caused by deletion of an item forces the reader to utilise other sources for the information needed. These will be available from immediate contexts, both unilateral and bilateral, contents of working memory and contents of other permanent stores especially of linguistic items.

Any deletion of words that requires the missing items to be supplied implies that after the concept, idea or relationship needed for meaningful completion has been selected, a search is activated via retrieval mechanisms for the specific verbal form that the reader judges will best fit the frame. The search process as an explanation for cloze is suggested by J. Jaap Tuinman (1972:75), and fits in very neatly with our model of the reading process. The input as comprehended will itself contribute to the on-going discourse and in turn, impose constraints on the selection of the subsequent missing items. The latter will be selected from items in permanent store or from those available in STM and WM because they form part of the language of input. The search process will be re-activated for each deletion, and knowledge of the language will enable the reader to predict the class or category that fits in with input, that is, 'sequential predictability' (Smith, 1971:20-21) will be made easier. Knowledge of the meaning of the discourse up to that point, that is, the contents of WM will help narrow down the range of options within the

class so that 'distributional predictability' (Ibid) is possible.

Where the deleted item comes from a very small class, the search will be faster; and where the class is larger, retrieval will take longer since more than one alternative is available, and the reader will have to choose the one he thinks provides the best fit. The size and range of items in store however will vary, as we made clear in our model, and what is stored will reflect the competence of the NNS. The classification and categorization of items in store will reflect the level of language learning and proficiency possessed.

To refer back to our model, we stated how attention and distribution of processing capacity reflected the level of language ability, since higher level processes could only occur if the demands at lower level were minimal. This capacity has now to be shared between processing of input, and processing and production of output responses. As long as the demands of both maintain a balance, the reader will be able to cope. This will be possible when the deleted item is easily predictable and can be retrieved from store with the help of existing contextual cues. As long as the item conforms to frequently used patterns, the reader will easily supply the missing text. But when the reader is unable to retrieve from memory a fit, the balance between distributed capacities will be upset. Because response

generation is the focus, input processes will have less attention, and higher level processes which are concerned with long range constraints will come to a halt. From the model, one can predict that semantic constraints at higher level would be the first to disappear followed by syntactic and grammatic constraints imposed by decreasing context. The implications here are that with increased item difficulty or poorer language ability, items will be output that violate the norms and rules of use and usage since the various constraints operating on MVD will have ceased to function. Specific hypotheses will be used to evaluate our model of reading which will then enable us to evaluate ability and comprehension of the NNS. These will be postulated in the next chapter. But what we have done is to attest that cloze as a technique has construct validity since it activates the same processes we have hypothesised are entailed in the reading activity.

5.3 SUMMARY

To summarize then, our model has attempted to synthesise known facts about the physiology of reading and their psychological implications. Based on theories and models reviewed in Chapter Four, our five-stage model has been primarily designed for written input and output. Components and processes have been specified functioning under limitations of the system, and it is hoped that our elaboration of both the process and the end-product which

is represented by the contents of working memory will enable us to produce testable hypotheses. It is assumed that an analysis of reader's responses in various tasks that reflect the model will reveal insights that will permit reliable and valid evaluation of reading efficiency.

CHAPTER SIX

DEVELOPMENT OF HYPOTHESES AND DISCUSSION OF TASK VARIABLES

6. GENERAL REMARKS

In our discussion of the reading model, we stated that the efficiency of the processing of printed matter would be dependent on two major factors:

(i) The first was the level of proficiency determined by the size and distribution of language components in permanent store. This enabled the reader to recognise, identify and assign meaning to units of input in immediate, short term and working memory i.e. variations in size and complexity of the items in store would reflect levels of proficiency.

(ii) The second factor was the utilization of cognitive skills and strategies that enabled the reader to reduce input and re-organize the results of earlier processing for efficient storage in response to the task requirements for which the reading was undertaken.

If one was interested in investigating the process of reading one needed to attempt evaluation of the comprehension while the input was undergoing conversion from visual form to meaning. And if the focus of interest lay in the end-product of the activity, one needed to delay assessment

until input had been completed. Since one of the major aims of this study was elicitation of data that would permit us to make inferences of the cognitive activity - which as our model had postulated was involved in reading - we used the cloze technique to evaluate the process stage. For the product, two techniques were utilized: the writing of a summary and the use of multiple-choice questions, both of which would require the reader to evaluate and re-organize the contents of working and long-term memory in response to specific task-objectives. The general hypotheses tested then were based on the reading model, while the specific hypotheses were dictated in part by the choice of techniques selected for the experiments.

6.1 DEVELOPMENT OF HYPOTHESES

6.1.1 Evaluation of the Reading Process

If as we suggested in Section 5.2.1., cloze necessitates the reconstruction of a partially mutilated text by reference to items in various memory stores, it follows that the results of the search process will reflect the quality and quantity of the linguistic components stored. And since the success of the search would be directly related to demands made by language in use, this implied that success would also be indicative of linguistic competence. Standardized test batteries like the ELBA and EPTB which aim at eliciting similar

evidence of proficiency in English as a second language, utilise a large number of discrete test items so that the resulting overall score provides reliable data by which the level of linguistic ability can be inferred. It follows that reliability in a cloze test also is necessary if claims regarding its ability to differentiate between levels of linguistic proficiency are to have any validity. Our first hypothesis then was to establish cloze as a valid measurement device.

Hypothesis 1 Tests of cloze requiring the successful completion of a large number of deleted items will correlate substantially with standardized tests of English as a second language, thereby establishing the degree of relationship between reading ability and linguistic proficiency in NNS.

This meant that subjects achieving higher success scores would be interpreted as possessing a higher degree of linguistic proficiency as compared to those obtaining a lower score, and this in turn would reflect reading ability. Referring back to our model, lexical stores too could be specifically evaluated in the search task that cloze represents. Where clearly defined boundaries existed, the search would produce items that reflected a degree of similarity to the deleted item. Even if not identical to the author's words, the responses would be similar in meaning and so acceptable as alternative

responses provided they conformed to constraints imposed by grammar and syntax. Therefore the reader with a wider vocabulary in the different categories of meaning and corresponding classes of word-forms, was more likely to have a greater choice of options than the reader whose vocabulary had not yet been classified by lexical fields or form classes. The latter was therefore more likely to have a wide choice of items that did not fit the constraints imposed by context. This meant that NNS of this type would have less conformity of responses. Since the alternative responses that were acceptable would benefit the more proficient reader, we expected to demonstrate this by using two statistics, the mean and the correlation coefficient.

Hypothesis 2 When cloze items are scored by verbatim i.e. exact word replacement and by acceptable word methods, the gains in scores will be substantially higher for the more proficient readers. Differences in scoring systems can alternatively be measured by correlation, with lower values obtaining for the better readers as compared with values obtained by the less proficient readers.

In Sections 2.4.3 and 5.2 we suggested that with length, the information would tend to become increasingly passage-specific, and the competent reader

would find it easier to predict the missing items aided by redundancy. With shorter texts, the difficulty was likely to be greater because of the greater density of information. However if the reader was not competent, the continuing demands made on the cognitive processes by a long text would cause the gap between writer and reader to widen leading to a breakdown in communication. We expected to confirm these arguments by testing the two following hypotheses.

Hypothesis 3 The longer the text, the easier it is to discriminate between different levels of reading ability i.e. between more and less proficient readers.

Hypothesis 4 The number of omissions representing failure to retrieve any word that might conceivably fit the constraints imposed by context, is likely to increase with length of text if the reader is not proficient in the language. This is likely to be accompanied by a reduction in speed of performance.

These four hypotheses were concerned with reading performance during input of printed matter and were necessarily focussed on cloze-type tasks. As is evident, we were interested in utilising for analysis the total response patterns of non-native speakers and therefore proposed a breakdown of responses into: success scores

scored by two methods, incorrect items, and omitted items. In fact we expected to obtain from cloze far more by way of value than merely viewing it as a measurement device. Weaver (1965) had suggested that the cloze technique afforded a close-up view of what was occurring at precise points in the passage, and albeit presenting a distorted view, revealed more about the process than other techniques. We planned to utilize experimental data to develop a reading skills profile based on the total response patterns that would help us to identify the 'good' from the 'poor' reader.

The fifth hypothesis was the central issue raised in our model and the reason for the selection of different techniques.

Hypothesis 5 Evaluation of the process - via cloze - though positively and significantly correlated with the evaluation of the product of reading, will have sufficient unexplained variance to confirm that they two separate aspects of the reading activity are being reflected.

By comparing cloze with two measures of reading product i.e. summary-writing and multiple-choice questions, we hoped to establish the validity of a reading model that viewed process and product as related through identifiable entities.

6.1.2 Evaluation of the Reading Product

As stated earlier we had selected the multiple-choice technique and a summary-type evaluation to measure the subject's understanding of the reading after input had been completed. These two techniques utilised the contents of working memory when they followed the reading. It was expected that deficiencies in language knowledge would be partially compensated for by skilled use of strategies, and although comprehension of the end-product depended to some extent on comprehension while processing was occurring, the two represented different aspects of the reader's ability. Since higher level processes related to larger linguistic units of discourse, they represented a more complex level of activity and would normally occur in working memory. Questions concerned with smaller units of language would be processed faster and more easily in short term memory, which led to two hypotheses.

Hypothesis 6 Questions of detail explicitly stated or implied will be easier for all readers than questions involving more generalized concepts.

Hypothesis 7 Items with distractors taken from a localised context will be easier than items where the distractors occur in widely separate contexts.

These then were the specific hypotheses that were tested in the experimental investigations. However since there is much more to a test than a measurement device, let us turn now to a consideration of the other factors involved.

6.2 DISCUSSION OF TASK VARIABLES

Just as our model was reader-oriented, and hence concerned primarily with the cognitive processes working on the input, in this section we need to discuss factors concerned with the printed materials that served as stimulus for the input. To repeat what was said in Section 2.1., because we know so little of the internal processes the only controls we can exercise, are generally related to the 'task-induced operations' (Frederiksen, 1972:211) we set up to elicit performance. By attempting to ensure the reliability and validity of the instruments or techniques we use to measure this performance, we hope that what is later inferred regarding reading ability is not too far from the truth.

The accompanying diagram in Figure 10 presents an outline of the three major factors we will discuss : the techniques selected for the measurement which by their very nature represented the task; the scoring procedures that were adopted for each of the techniques and tasks; and thirdly the materials used in the operation.

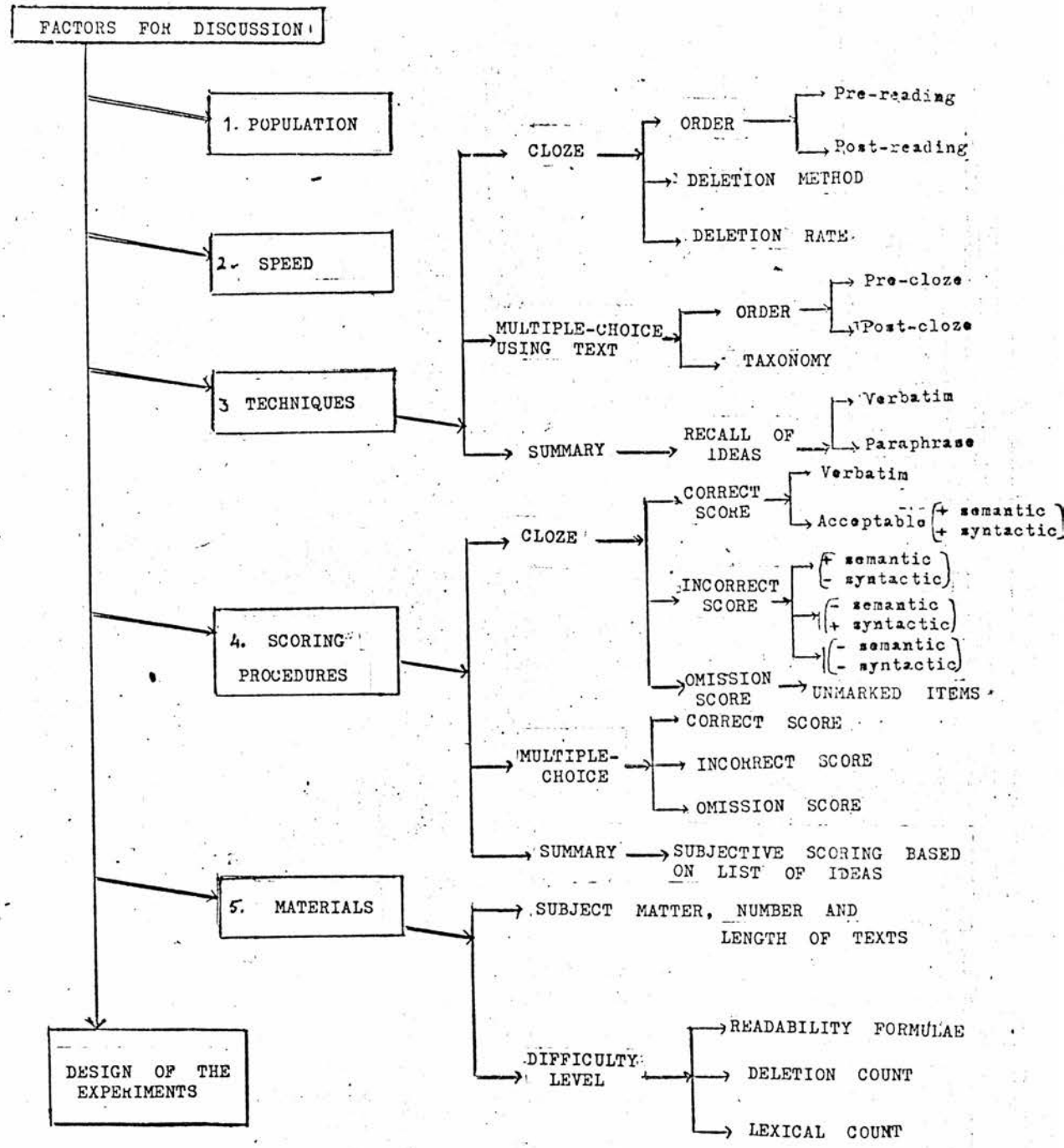


Fig. 10 TASK VARIABLES TO BE TAKEN INTO ACCOUNT

But before we begin the discussion, a brief word about the two other factors not covered by the three major variables will be useful.

6.2.1 Population

Because our population sample was fairly clearly defined we made two assumptions that were justified by the context.

(i) NNS wishing to enroll in courses of instruction conducted in a foreign language, either at home or abroad, would be relatively uniform in intelligence as compared with all other possible subjects who had stopped their education at an earlier level.

(ii) NNS would be highly motivated in a foreign language testing situation, realising that success in their studies would be affected by their ability to both read and perform well in English.

The population samples used in the experiment were taken from two major sources. NNS who had gained admittance to a British university and who had either already matriculated or planned to matriculate, formed the bulk of the subjects in our investigation. Overseas samples came from two countries, W. Germany and Egypt. Both these groups comprised local students who were receiving instruction in English because of the requirements of the specific courses they had enrolled on. They were

not attending one foreign language class among other courses conducted in the first language, the entire courses were in English. To this extent they resembled the group in Britain.

Further details will be given when the experiments are discussed in Chapters Seven and Eight.

6.2.2 Speed in Reading

This is an aspect that we feel deserves more attention since reading proficiency alone is not enough. It needs to be combined with speed so that the resulting performance is a combined measure of both 'speed' and 'power', or as we preferred to call it 'reading efficiency'. For us the latter was a measure of maximum gain in minimum time.

Miles Tinker reported as far back as 1939 that speed versus comprehension in reading were affected by the level of difficulty of the material so that the two measures correlated more closely when measured on identical material, less so when compared on different materials. In a discussion on the same subject Francis Robinson (1940) raised an important issue which we too had seriously considered. When material read was easy, more questions were attempted resulting in higher scores. But when the material read was difficult, one generally assumed that unmarked questions represented unattempted questions. But how could one assume this?

Perhaps the reader had considered the question and failing to find a satisfactory answer had left the question unmarked, giving the impression that he had skipped it without really trying. Tinker (1940:559) in reply stated that 'the only adequate definition of speed of reading is to consider it rate of comprehension'.

For these reasons we imposed a time limit in all the experiments that we judged was sufficient for all items to be attempted. All unmarked items were classified as 'omissions' and represented failure due either to inability or speed.

6.2.3 Testing Techniques

Carroll makes a very useful distinction between comprehension ability and comprehension of a specific passage. It was the former that we were interested in, realising that it would have to be inferred from the second which was what would be directly evaluated. Discussing the suitability of alternative techniques Carroll suggested the use of 'good multiple-choice, true/false, evaluation of main ideas requiring construction and re-organization, following verbal directions ... and techniques that depended on context and redundancy like standard cloze...' (1971:89) as measures of general comprehension ability. For the specific comprehension measure he felt that all the above, except perhaps cloze, were suitable including recall of a paraphrase of the passage which was a valid test, but rarely used because

of the difficulty of formulating adequate scoring procedures. He voiced uncertainty about the value of cloze here stating that there was a lack of clear evidence about the ability of cloze to measure comprehension of major ideas or concepts that ran through longer discourse.

Our decision to use cloze for the major part of our experiments was made on the strength of our model, and reflected our interest in the process of reading.

Multiple-choice as a measure of ability and specific understanding of the text read reflected the end-product, as did the use of a summary. Both these latter techniques are commonly used in classrooms, and the multiple-choice format is similar to the reading comprehension sub-test in proficiency batteries (See Section 3.2. and 3.3). The summary however demands much more from the testee by way of productive evidence of comprehension. It also demands much more from the tester since subjective evaluation of the reader's performance must be made. For both these reasons its use was limited to two small groups in one experiment.

Each of the three techniques will be discussed in greater detail now.

6.2.3.1 Cloze

If 'a foundation of all language skills is the capacity to anticipate elements in sequence' Oller (1971),

then cloze would appear to have higher claims than most other measures as a technique for evaluating language proficiency. And 'a test requiring a subject to supply verbal material to fit a context - where the "correctness" of the response depends on the likelihood of the response in that context - has been found to be a good measure of the examinee's knowledge of his native language', Carroll, Carton and Wilds (1959:6).

Cloze has been defined by Taylor (1953:416) who first developed the measure into a technique for assessing readability as 'any single occurrence of a successful attempt to reproduce accurately a part deleted from a message (any language product) by deciding, from the context that remains, what the missing part should be'. The cloze procedure consists of a set of rules in accordance with which 'words of a passage are systematically deleted in some mechanical way, and replaced by blanks, usually of a standard length' Anderson (1970:178). The subject's task is to replace the missing words, and his score, the number of words he correctly replaces, is an index of his comprehension of the passage since it represents how difficult the task is for the individual reader.

The rate of deletion used in the cloze tests has ranged from every fourth (or even every second in some experiments) to every twenty-fourth word, and both Taylor and Carroll have reported that different forms of

cloze tests on the same passage, even with the same deletion rate, revealed different difficulty levels, depending on the words that had been deleted. However Bormuth (1969a) reported that variations in means and standard deviations all steadied with an increase in length, although the range continued to show wide differences in minimum and maximum scores. Research regarding rate of deletion had suggested that 80% to 90% of the original passage was best left available, that is, from every fifth to tenth word could be deleted without radically interfering with communication. Although Taylor favoured every fifth to seventh, Carroll et al in their 1959 study with NNS had used every tenth word deletion to give their subjects the advantage of increased context. For our experiments we, too, adopted a similar deletion rate, omitting every tenth word and replacing it with an underlined blank of fourteen spaces in length. As is normally done in most cloze tasks, the first and last sentences in each test were left intact.

Most of the literature on cloze reports experiments with passages of 150 to 250 words, or when longer texts are used, sub-samples are employed for the cloze tests.

We chose nine texts for use in the cloze experiments, five long and four short. All five long passages permitted 100 or more deletions, while the

shorter texts had 25 deletions each, all at the same rate with every tenth word deleted.

The random method deletion is accepted as being as good as any. Carroll (1971) had argued that deletion of lexical items only, implied that they alone carried the information load. But when functional words had been deleted, the level of difficulty had risen suggesting that some of the meaning burden was borne by them. They too must be involved in the complex interaction of language forms by which meaning is conveyed in English. Because it is almost impossible to define which linguistic forms are contributing to meaningful verbal discourse, the random 'any' word deletion was considered the most suitable. The ability to make use of redundancy features has been described by Weaver and Kingston (1963:259) as 'redundancy utilization', and it follows that when the grammar or meaning of a unit constrains the occurrence of other items in their neighbourhood, this is the principle at work.

This is perhaps better expressed by Gleason (1965:458-459) who suggested that the grammar of a language could be considered a set of restrictions on the occurrence of possible structures within the language. All languages contained elements of redundancy which operated at levels as varied as spelling, grammar, syntax and meaning. Cloze by eliminating words at intervals reduced the original text. But, provided enough of the context was available, it should be possible to recover

by inference from the remaining portion what the missing items were likely to be.

'If the encoder producing a message and the decoder receiving it happen to have highly similar semantic and grammatical habit systems, the decoder ought to be able to predict or anticipate what the encoder will produce at each moment with considerable accuracy. In other words ... they should be able to anticipate each other's verbalisations' Wilson and Carroll (1954:111). This is very similar to Cicourel's reciprocity of perspectives (See 2.2.7.) Cloze 'taps simultaneously all of the complex determinants affecting word choice both at various levels of organization and through long stretches of sequencing' Wilson and Carroll (1954:112). The technique then provided us with an objective measure of the language correspondence between writer and reader, and seemed eminently suited as an assessment procedure for NNS planning to pursue higher level studies in the foreign language.

Morecver because of the mechanical objective nature of the deletion process as we adopted it, the discourse presented was a sample of the author's language unaffected by intrusions in the form of questions. The relationship between writer and reader appeared more direct and natural than what was encountered in conventional techniques used to test comprehension ability.

Cloze may be administered to subjects who either have or have not seen the original unmutilated text.

Bormuth (1968) found there was little difference between cloze tests administered before or after the passage had been read, although as expected, scores for the latter procedure were higher and results correlated marginally better with other conventional tests of comprehension. But he and Rankin (1970) both agreed that the gain in saving time by using pre-reading cloze was more valuable, and suggested that by lengthening the test, exact scores might go up, contributing to gains both in correlation and reliability. The implications were clear. The longer the test the easier it ought to be resulting in higher scores.

Our reasons for using pre-reading cloze were primarily dictated by our interest in the processing of input at first exposure to the text. There was however one small experiment where pre-reading and post-reading cloze were attempted, permitting comparison of results.

We had then taken into account, rate of deletion, method of deletion and ordering of the cloze test. What remained was an analysis of the items deleted which would contribute to the ease or difficulty of the task. This is discussed in greater detail in Section 6.2.5.2 (2).

An objection to the use of cloze as a comprehension measure referred to in 6.2.3 needs elaboration. MacGinitie (1961) had conducted a series of experiments to find out whether constraints upon

words in continuous discourse were cumulative over more extended lengths of prose text. By varying deletion rates, he reported that 'results showed that additional uninterrupted context beyond five words did not help in the restoration of the missing words'. He however went on to say that this did not mean that no constraints operated beyond the four or five word distance. 'Some constraints, such as knowing the topic of the paragraph, may have a more generalized influence' (1961:128).

Carroll, Carton and Wilds (1959) had conducted an experiment aimed at this specific question, that is, whether performance of subjects was affected by paragraph cues. And the authors concluded that there was a general tendency for content words to be more affected by paragraph cues and wider contexts, than functional words. 'Scores on 20 unrelated items were very different from 20 items of a continuous passage', and diminished context decreased scores (Ibid: 110-111).

We utilised the cloze technique for nine texts, using the 'any' word deletion method for every tenth word. And to investigate the influence of context, conducted an informal experiment very similar to that by Carroll et al. This will be reported in Chapter Seven. We also adopted Rankin's advice (1970:136) that the concurrent validity of the cloze procedure as a measure of specific comprehension could be determined

by correlating cloze tests with scores on comprehension tests covering the same material. We used the multiple-choice technique to do this.

6.2.3.2 Multiple-Choice

With this technique there is no danger of being restricted by the task to a limited context. One of the major advantages of this technique is that it is open to manipulation so that the tester can vary both the level and aim of the specific questions designed to test comprehension to suit his own purposes. Obviously if not treated with care, this advantage can also become a source of interference.

Not only do cloze and multiple-choice tap two stages of the cognitive processing system, but the techniques themselves complement each other in a number of ways, each providing useful information. In our discussion of the reading model we stated that comprehension was realized by the end of stage three of the process. Where recognition of the correct answer - from a range of alternatives already provided - was sufficient proof of the reader's ability to perform the task set, the extent to which productive expression had to mediate between comprehension and measurement, was reduced.

However the injection of questions between the writer's text and the elicitation of response markers

from the reader, necessitated the understanding not just of the original discourse, but also the problems set by the use of this technique. And an understanding of the former was inferred by the skill exhibited in the latter. The communication of the message from writer to reader was now assessed through the performance of tasks based on what the tester assumed was the information core.

In a technique developed primarily as an objective and reliable assessment measure, the index of ability was paradoxically based on a far more subjective criterion than is generally admitted. The approximation the reader achieved to the tester's understanding of the text would enable him to score well in the tasks set. The value then of a good multiple-choice test to do what it sets out to do is in reality a test of the examiner's ability to use the technique efficiently. The efficiency of a cloze test depends far more on the materials used and the ease or difficulty the individual reader has in completing the deleted items; whereas the value of a multiple-choice test is more dependent on the questions asked and the reader's ability to answer them than on the text itself. Both measures however have construct validity. Both need the exercise of similar skills for many of the tasks, but whereas for cloze one assumes their presence because of logical necessity, with multiple-choice the specific

question set can focus attention on either inferential processes, or evaluative or organizational etc. It is this latter ability that gives validity of content in the test, especially when reading must be undertaken at this advanced level.

According to Wilga Rivers (1968:294) 'Reading comprehension and reading comprehension alone, is more validly tested by selection of the correct answer from multiple-choice items in a foreign language, on a reading passage. In this case the student shows his ability to comprehend not only the reading passage but also the fine distinctions between the choices offered him'. Since the twenties,[1] attempts were made to present comprehension ability as a set of discrete, definable skills, which if learnt would lead to understanding, or more strictly would result in understanding. Burkart (1945) listed 214 separate abilities which her survey of the literature had produced. Davis (1944, 1968, 1972) reduced his original list of several hundred to eight which he claimed he had isolated. But attempts to replicate his findings reduced the range to a probable five. Other educationists like Bloom (1956,1971) approached the task of attempting to define the ability by presenting a taxonomy of educational objectives that concentrated on the 'cognitive domain' but which did not exclude factors from the

---

1. F.B. Davis has an excellent review of the earlier literature (1971).

'affective domain' that might influence reader participation. These critical skills were operationally defined, and sample tasks included for exemplification. Bloom placed comprehension at one stage higher than knowledge which he said 'could be mere rote learning or verbalisation' (1971:149-150). Comprehension was defined at three levels, (i) translation tasks like reading musical scores, interpreting maps, pictures etc. (ii) interpretation of verbal matter, and (iii) extrapolation which involved application of knowledge as evidence of comprehension.

For our test using the multiple-choice technique, we drew up a taxonomy of objectives that enabled us to produce items for inclusion. Although these were then subjected to the usual statistical procedures of item analysis, when items had to be replaced or revised, the existence of the taxonomy was a useful indicator guiding us in our decision. Some of the items then were in use with larger numbers than others of the population sample. All items were based on the same long text, so that we had an experimental design that permitted us to compare two techniques, cloze and multiple-choice, based on the same passage.

The taxonomy was designed to test whether NNS possessed the following abilities necessary when reading for information:

(i) ability to extract factual information, both when distractors are taken from localised context and when taken from different parts of the text.

(ii) ability to recognise referential devices like anaphora, substitution etc. that writers use to give cohesion and coherence to discourse.

(iii) ability to make inferences regarding information in the text using both localised context and when wider contexts are involved.

(iv) ability to evaluate information either in the passage itself or by reference to the passage, such as the best summary or determining fact from opinion etc.

(v) ability to recognise or infer main theme(s)

(vi) "other skills" such as ability to define, or use organization of text for re-ordering contents in some specified manner e.g. sequencing etc.

Because of the greater demands imposed by the technique, this format was only used with one text, but formed part of the main experiment. Item cues avoided external evidence of length, or syntactic and grammatical disagreement in distinguishing the correct option from the alternatives, so that the choices attracted some degree of attention from the less skilled readers. Like many multiple-choice comprehension tests, we offered a one in four choice, with the correct answer randomly distributed among the distractors. Distribution of

the questions according to the taxonomy is given in Appendix 1B.1.5 (page 27)

For this part of the experiment, the burden on LTM was partially alleviated by having NNS read the text and answer the questions with the text available for reference. By presenting text and questions together with a combined time limit, it was hoped that NNS would adopt whatever strategies they wished to enable them to process the input efficiently. And in the main experiment the ordering observed was: pre-reading cloze, followed by the multiple-choice test comprising the same text in unmutilated form accompanied by questions. We did however conduct one experiment where the two tests were presented in counter-balanced order. In another sample NNS did only the multiple-choice test permitting comparison of means done after and without the cloze test.

It has often been suggested especially in the recent literature (Pyrczak 1975, Tuinman 1973) that it is often possible for subjects to answer multiple-choice questions even without reading the text so that questions should be tried out prior to use in an experiment. But these tests are aimed primarily at information gain as a result of reading a specific passage. Our concern was that given a level of textual difficulty, NNS should be capable of abstracting the information necessary for specific search questions.

We had moreover found as a result of an informal experiment with the cloze technique (to be reported in Chapter Seven) that readers are adept at imposing form and coherence even when the text was fragmented and comprised a series of apparently unrelated chunks. Hence no attempt to pre-test questions without the text was made.

To sum up, the production of test items using this technique was aided by use of a taxonomy of abilities that NNS were expected to need to read efficiently. The items demanding the recognition of the best option, offered a one in four choice. And the ordering adopted for the multiple-choice test in the main experiment, was as the second part of a two-part test, the first being a pre-reading cloze task using the same text. The unmutilated text was always presented with the multiple-choice questions.

6.2.3.3 Summary-Writing

This third technique involved a recall of the text read since this was not available for use during this part of the experiment, and in terms of our model utilised the contents of WM. The major difference between the recall and the multiple-choice technique was that the former required the reader to re-encode all the contents of WM that he considered essential and which best reflected the meaning of the discourse as a whole.

The selection of information was left to the reader but ability to comprehend had to be matched by ability to re-encode in English. Like the cloze technique, this recall by paraphrase task involved the full five stages of the reading model.

Because recall of the gist simulates the normal reading situation we were not interested in verbatim recall. The difficulty in using this technique lay in the fact that we used it as a measuring instrument. We used a list of ideas that we pre-determined adequately represented the information content, but the problem of matching these with the readers' verbalisations remained. Because of the subjectivity of the scoring, this technique was not extensively used, being employed with two minor samples only.

In each experiment, subjects were told to expect a summary question, and in both cases it followed immediately after subjects had seen and read the passage. In one experiment, it followed immediately after pre-reading cloze, and in the second an unmutilated passage was presented for reading after which the paraphrase by recall was required.

These then were the three techniques utilised. We will now discuss the problems of scoring each of these.

6.2.4 Scoring Procedures

Since evaluation of reading ability was in terms

of scores, the various procedures including statistical measures that were to be used for the scoring were obviously an important factor to be taken into account. The literature on all three of the techniques we used presents a variety of scoring procedures that have been tried out with varying degrees of success. But as we were not interested in the techniques per se, but in how well it evaluated what we wished to test, we used procedures that had been established by practice. In addition we evolved analytic approaches that permitted us to produce data that could be applied to the hypotheses being tested.

6.2.4.1 Cloze

The two most widely used scoring systems have been the exact word and the acceptable word methods, although the latter has been interpreted in a variety of ways. For the exact word method the 'score of any subject on any passage is the number of his proposed words that match the original ones deleted, and any passage's score is the total of the scores all subjects make on it' Taylor (1957:20). The first gives us the individual's reading success score while the second helps determine the readability or difficulty level of the passage for the population taking the test. The exact word method has all the advantages of reliability and objectivity. And of course, what is invaluable when defining task

objectives concerned with language proficiency, it provides a clearly defined criterion measure which is so difficult to identify, let alone specify in most measuring instruments. We used this method, counting verbatim replacements and ignoring errors of spelling and punctuation provided the word was clearly decipherable and grammatically in concord with the surrounding context. However as our discussion of the cloze procedure indicated, provided comprehension had occurred and the NNS had sufficient knowledge of English to enable him to verbalise his understanding by predicting the missing item, the chances were high that he would produce a response that was very similar in form and meaning, if not identical, to the author's word. Sequential and distributional predictability would be guiding his choice of word. Also because, as Oller said (1973:109), there was something counter-intuitive about requiring the exact word when dealing with non-native speakers, all results were assessed in two ways.

As our diagram indicates, 'acceptable' for us constituted any item that satisfied the semantic and grammatical constraints imposed by the context. In many studies, the latter criterion alone was considered sufficient for a response to be judged 'acceptable', and words of the same form class were scored, Fillenbaum, Jones and Rapoport (1963).

But since our study was concerned with

comprehension, we preferred to use evidence that reflected sensitivity to both form and meaning. There was no doubt that introduction of acceptability criteria affected the objectivity of the cloze procedure. On the other hand, had it not been adopted, verbatim scoring would have imposed an unnecessary dichotomy into the assessment forcing an item to be scored as 'incorrect' because of the author's preference of one word. It would have meant refusing to give credit even when comprehension had obviously occurred. So though cumbersome, both systems of scoring were utilised, and a hypothesis set up to estimate the value of the alternative systems. Once again spelling and punctuation were ignored provided the item could be unambiguously interpreted. Where corrections were made, or two alternatives provided, the last word in the blank space was scored, and the rest recorded but ignored in assessment. Multiple word responses were counted as incorrect since instructions asked for one word only per blank space.

Carroll (1971:81) had suggested that verbatim or 'strict' scoring might represent general ability, and the more 'relaxed' scoring procedures represent specific comprehension of the passage used.

The question in our study was what procedure to adopt for determining 'acceptability'. We fell back on established practice, and for each of the nine passages used for cloze, three to five native speakers

were asked to supply the missing items. Their responses formed the basis of a list of acceptable alternatives for each item. However for the text used in the main experiment, we submitted the results to a panel of three native speakers. These three judges were given the pre-reading cloze passage along with the list of replacement items. The original exact word, where missing from the alternatives, was inserted in random order into each set of choices. The judges were instructed to select all words that in their view would be acceptable in the context, listing them where possible in order of priority. Each judge worked independently, their task being to choose multiple alternatives from the multiple choices offered. They were also asked to supply items which were not present but which in their opinion would be 'acceptable'.

As expected there were some instances in the content category where judges disagreed. A second arbitration panel of two native speakers was used to resolve this difficulty. Eventually a list of acceptable items was prepared. During the experiment, this list was marginally expanded to include alternatives that in the opinion of this writer conformed to all the constraints imposed in the preparation of the earlier list.

All responses then were scored for:

(i) correctness, that is, acceptable responses (including exact word responses).

and (ii) incorrect responses.

Omissions were also scored and recorded since one of our hypotheses was specifically aimed at evaluating the extent to which failure to respond reflected reading disability for the NNS. Each correct item would carry the value of one. Total scores possible were 25 for the shorter tests, 100 for four of the longer texts, and 115 for the fifth test.

6.2.4.2 Multiple-Choice

Like cloze, the scoring for this technique entailed a record of correct, incorrect and omitted items. The correct score was the number of items where the option selected matched the one pre-determined by the author as the correct response. The selection of items had however, in most cases, been determined by item analysis so that 'correctness' reflected the opinion of more proficient readers.

The total possible score was 20 for the multiple-choice test except with a small sample who attempted an additional 5 items on the same text. There were three experimental versions used.

Item analysis techniques were also used for the cloze tests but as descriptive measures only. No item was revised or deleted from evaluation because it

failed to discriminate between good and poor readers, or because it proved too easy for all. The data from cloze was treated as "true scores" in statistical analysis in conformance with an opinion expressed by Lee J Cronbach and quoted by Taylor (1957:20) in a footnote that 'cloze results appear to satisfy the assumptions for scores, but not those for frequencies because successive blanks may not be independent. Semantic and structural factors tend to make for interdependencies among the words in a meaningful series such as an ordinary sentence.'

Reliability measures were carried out for all tests, and correlation coefficient of two sample groups on cloze and multiple-choice with two standardized batteries, the ELBA and EPTB already referred to, provided the necessary validity using an external criterion.

6.2.4.3 Summary-Writing

As stated earlier, the difficulty of obtaining an objective evaluation prevented the use of this technique on a larger scale. David King (1960) had tried out seven different schemes in an attempt to see which best reflected the accuracy of written recall, and he concluded that a list of 'ideas' was one of the best measures, with length of response as a second though equally vaguely defined indicator. To measure 'ideas' he counted exact words present in the recall that also

occurred in the original; he also used synonyms and cloze type tests, and concluded that the most reliable indicator appeared to be content words. This approach however seemed too indirect as a measure of comprehension and we fell back on a list of ideas produced in two ways. The first was to ask four native speakers to write a summary using the text to aid them. The second method was by performing a crude rhetorical analysis of the two texts to be used in the experiment. This analysis was based on the flow diagram technique reported by L.F.Thomas in an undated mimeograph. Details of how the procedure was adapted for use are contained in Appendix 3. It was decided to score each idea independently. Out of a possible total of 25, the individual's score was the number of ideas or propositions that matched with the list. In the instructions to the summary, NNS were asked for the main points only, in note-form if they preferred. Because of the grammatical inaccuracy of many of the responses, a subjective decision had to be made regarding equivalence in meaning between the NNS' writing and the criterion list. Only correct scores were reported for this technique and results were correlated with performance on cloze tests. With one sample the cloze test was based on the same passage and preceded the summary; with another sample the cloze test was based on an independent passage.

To summarise then : three techniques along with their own scoring systems were described. The success or failure at the task-induced operations was used to infer the reading competence of the subjects. Five major variables involved in the testing situation were identified for discussion. Of these four have already been dealt with. This brings us to the fifth major variable, the linguistic materials used in the entire testing operation.

6.2.5 Selection of Materials

Because this is an area over which the experimenter has the most control, it inevitably received a great deal of attention. Our choice of materials, though subjective, was determined by two major considerations.

(i) Firstly, the texts had to be self-contained units on non-specialist subjects and of a specified length.

(ii) Secondly, these texts had to be roughly comparable in level of difficulty as far as could be objectively determined. The latter problem was solved by using the reading comprehension sub-test of the ELBA (Ingram et al) as a criterion measure.

6.2.5.1 Subject-matter, number and length of passages

Nine texts were selected from a variety of sources to permit us to compare techniques and variations in length of material read. Five of these were over 1000 words in length permitting a deletion count of 100 or more items for the cloze tests. The remaining four passages were shorter, varying from 280 to 340 words permitting a deletion count of 25 items each in the cloze.

One of the five which was used in the main experiment was about the development of civilisation and will be referred to hereafter as Passage 1 or P1. The other eight formed four pairs of texts, each pair comprising one long and one short text taken from the same source, written by the same author and on the same topic. Some editing had to be carried out to ensure that each text was self-contained, but changes were kept to the minimum necessary.

These eight texts which were used in one experiment were all presented with titles, the long and short of each pair carrying the same title. They will when necessary be referred to by their numbers or their titles. For passage 1, a title is provided for convenience.

Title	Length	Source
Passage 1 Civilisation	1096 words	The Human Species by Anthony Barnett, Penguin (1950;1968) pp.216-221 (excluding 218 which was a table)
Passage 2 (A):P2A What is Law?	1063 words	What is Law? by Ronald Rubenstein in Topics and Opinions (1956) Editor: A.F.Scott London: Macmillan & Sons.
Passage 2 (B):P2B What is Law?	287 words	-ditto-
Passage 3 (A):P3A Urbanisation	1051 words	English in Social Studies: Unit 7 Editors: J.P.B.Allen & H.G. Widdowson. OUP (forthcoming)
Passage 3 (B):P3B Urbanisation	306 words	-ditto-
Passage 4 (A):P4A The Leatherback Turtle	1203 words	The world's largest sea turtles remain a reptilian mystery by J.S.Bainbridge Jr. in Smithsonian September 1974
Passage 4 (B):P4B The Leatherback Turtle	320 words	-ditto-
Passage 5 (A):P5A The Scientific Study of Social Development	1071 words	A Philosophy for a Modern Man by H. Levy (1938) London: Victor Golancz
Passage 5 (B):P5B The Scientific Study of Social Development.	333 words	-ditto-

TABLE 1. SOURCES, NUMBER AND LENGTH OF PASSAGES SELECTED

Copies of all materials used in the experiments are contained in Appendix 1.

6.2.5.2 Difficulty Level

This problem was not so easily solved since obviously the level of difficulty of written matter depends not just on the linguistic complexity of the material but what I have called 'information density'.

Jeanne Chall writing in 1958 suggested that there were four elements that were significantly related to reading difficulty :

(i) Vocabulary load, which could be measured either in terms of diversity or difficulty, the latter determined perhaps by low frequency of occurrence.

(ii) Sentence structure, where the commonest measures were length, percentage of simple versus complex sentences, or perhaps percentage of sentences having prepositional phrases etc.

(iii) Idea density, 'From a logical standpoint we may assume that difficulty is related to density of ideas. We should be able to find that materials which are hard to understand contain more ideas to a given number of words. This element, however, has been extremely difficult to measure' Chall (1958:46).

(iv) Human interest or directness of approach This is supposedly connected with the relative number of personal pronouns and personal words like proper names, words like father, mother etc.

6.2.5.2 (1) Readability Formulae

Most researchers who have had to use readability formulae have reported that there are two variables which are fairly good indices of difficulty, Bormuth (1969b) and Klare (1974). These are some measure of semantic difficulty generally reflected in the vocabulary, and some measure of syntactic difficulty often reflected in sentence length.

From our point of view, we were more concerned with a measurement device that would provide the confirmation or otherwise that passages selected were of a more or less similar level for purposes of the experiments. We were not really concerned with the predictive value of these measures.

None of the formulae available measured all the four areas outlined by Chall, so we decided to use two that did not seem to overlap : the Fog Index developed by Gunning (1952), and the MTUL or Minimal Terminable Unit Length, proposed by Hunt (1966) exclusively for gauging syntactic complexity. Klare (1974:97) had suggested that 'sentence complexity is probably the real causal factor in difficulty, but length correlates very highly with complexity and is much easier to count'. But it was as well to keep in mind that when used in a readability measure that was a predictive device, this factor was only an index, not the cause.

Although the two measures selected graded the

four ELBA passages in the same rank order, they did not apply as neatly to the passages we had selected. The accompanying table will make this clear.

Passage		MTUL		FOG INDEX		CORRELATION [1]
ELBA 1	(80 words)	20	4	18.16	4	
2	(104 words)	13	2	15	2	
3	(152 words)	19	3	15·3	3	
4	(99 words)	12·25	1	11	1	
Average for ELBA		16·06		14·86		r = 1·00
Passage 1		20·68	1	15.57	5	
Passage 2	A (Long)	22·15	3	14·19	2·5	
	B (Short)	22	6	13·99	1	
Passage 3	A (Long)	21·44	4	16·5	7	
	B (Short)	23·5	8	17·9	8	
Passage 4	A (Long)	21·48	5	14·17	2·5	
	B (Short)	22·86	7	14·52	4	
Passage 5	A (Long)	21	2	15·87	6	
	B (Short)	23·78	9	16·97	9	
						r = ·3

TABLE 2 DIFFICULTY LEVEL (1 Easy to 9 Difficult)

The MTUL is a measure of the average number of words in a main clause, while Gunning's Fog index takes sentence length and the number of difficult words defined by length into account. Words of three or more syllables are counted as long, and the result is expressed in American

1. Spearman coefficient of correlation, Lewis (1972:58)

grade levels, Grade 13 being equivalent to undergraduate and 17+ to postgraduate. Unfortunately, Hunt has no similar grading device, but the subjective opinion of native and non-native teachers put the passages at academic level.

Later, a further set of readability measures was applied to enable us to see how they correlated with each other, and where they differed [Details of formulae are given in Appendix 2].

The Fry formula (1968)[1] was used mainly because it is the only formula that was developed when English was being taught to non-native speakers. Used in Uganda, it uses three samples of 100 words from the beginning, middle and end of the passage and uses sentence length and word length - the latter arrived at by a count of the total number of syllables in the sample.

The SMOG (Simple Measure of Gobbledygook) developed by McLaughlin (1969)[2] uses ten sentences from three samples taken from the beginning, middle and end of the text. The words with three or more syllables from the thirty-sentence sample alone are used, and this makes it the easiest formula to apply, and as we will see as reliable as some others. This formula could not be applied to the four shorter passages since their total count did not exceed 14 sentences each.

---

1. See Gilliland (1975:99-100)
2. See Gilliland (1975:94)

One of the oldest formulae, the Flesch Reading Ease developed in 1948, and which claims to measure level of abstraction turned out to be too complicated to apply, so the Farr-Jenkins-Paterson (1951)[1] formula that is considered equally predictive of reading ease was selected instead.

This formula employs a count based on the number of one-syllable words and sentence length, and results are interpreted from 100 easy to 0 difficult. A score of 30 to 50 indicates a difficulty level associated with academic work.

The sixth formula to be used was chosen specifically because it used the cloze technique as a readability measurement device, and it was thought it would be interesting to see how the actual results compared with the Bormuth formula (1969)[2]

The results of all the formulae are presented together for convenience rounded to one decimal point, and along with the rank ordering of these passages for comparison.

---

1. See Klare (1974:72)
2. See Klare (1974:80-81)

FORMULAE	P1		2A		2B		3A		3B		4A		4B		5A		5B	
FOG	15·6	5	14·2	2·5	14	1	16·5	7	17·9	8	14·2	2·5	14·5	4	15·9	6	17	9
MTUL	20·7	1	22.1	3	22	6	21·4	4	23·5	8	21·5	5	22·9	7	21	2	23·8	9
FRY	College		Gr.12		Gr.12		College		College		Gr.12		College		College		College	
SMOG	13·4	3	12	2	-		14	4	-		11·6	1	-		14·4	5	-	
FARR R.E	44	3·5	53	1	44	3·5	43	5	34·9	9	45·7	2	41	7	42	6	39	8
BORMUTH	29·7	4	32	1	31·7	3	22·8	6	22·5	7	29·6	5	21	9	22·4	8	31·9	2

TABLE 3 DIFFICULTY LEVEL : READABILITY FORMULAE

Rank Order: 1(Easy) to 9(Difficult)

Since the FRY and SMOG formulae could not be applied to the shorter passages, the remaining four formulae were used to provide the combined ranks of all nine texts. The oldest formulae, the Farr et al Reading Ease proved to be the most reliable as compared to the combined prediction. The correlation between the combined ranks and the Farr was r = ·95 using Spearman coefficient of correlation.

Combined Ranks:	Easy	Farr et al Reading Ease
2A		2A
1 & 2B		4A
4A	to	1 & 2B
3A & 5A		3A
4B		5A
5B		4B
3B		5B
	Difficult	3B

[Details regarding the formulae are contained in Appendix 2]

6.2.5.2 (2) Deletion Count

This particular measure was used since it was relevant to the cloze technique. We had suggested earlier that the difficulty level of a passage was directly related to the ease or difficulty of replacing the missing items. Using the 'any' method of deletion enabled us to get a sampling of items coming from the two

major categories often described as 'content' and 'functional'. In accordance with the usual practice we classified all nouns, main verbs, adjectives and adverbs of the -ly type as content; while all other words comprising deictics, auxiliaries, pronouns, prepositions, conjunctions and other adverbs not included earlier, were categorised as functional.

The total number of content versus functional items is given below, and as our survey of the literature had suggested, with native speakers it had been found that functional items were easier to replace than content.

DELETED ITEMS	1	2A	2B	3A	3B	4A	4B	5A	5B
CONTENT	49C	51C	12C	63C	14C	60C	17C	54C	15C
FUNCTIONAL	51F	49F	13F	37F	11F	55F	8F	46F	10F
TOTAL	100	100	25	100	25	115	25	100	25

TABLE 4 : DIFFICULTY LEVEL : CONTENT VERSUS FUNCTIONAL ITEMS

It will be recalled that all short texts had 25 deletions, and long texts 100, except P4A with 115 items.

Since content items were reported to be difficult, we used them to rank the texts from easy to difficult. For ease of comparison, text 4A items were reduced to a percentage.

Long	Easy	Short
1		-
2A		2B
	to	
4A		3B
5A		5B
3A	Difficult	4B

Since this difficulty count was directly related to the cloze deleted items, it was felt that this method may well turn out to be the most reliable predictor in the tests.

6.2.5.2 (3) Lexical Count

We come now to our final measure of difficulty level, and one that in some ways provided us with some very interesting information. This count was included because it seemed the only measure available of information density unaffected by sentence length or syntactic complexity. It was also a measure suggested by Frase and Rothkopf (See Section 3.5.2.)

Both Perfetti (1969) and Ure (1971) had suggested the same method of evaluating the content load by calculating the L.D. or lexical density of the text. This measure would provide an indication in percent of the content words as compared to the total orthographic count. As for our earlier analysis [See Section 6.2.5.2(2)] these content items included nouns, verbs, adjectives and adverbs of the -ly class which were assumed to carry the meaning load. While this is arguable, L.D. did provide an objective measure.

Ure reporting an analysis of spoken and written materials found that there was an average of less than 40% L.D. in the former. With written texts however, the lexical density was over 40%. 'Narrative has a higher density [than texts concerned with consultation, instructions, manuals, directions etc] and texts which are imparting information, whether narrative or other types of exposition have some of the highest densities of all' Ure (1971:450). Her higher density counts ranged from 47% to 58%.

Our analysis revealed the very interesting information that all nine passages selected for use in the experiments had a lexical density ranging from 44% to 56%, an extremely narrow range which provides support for Ure's findings.

The second measure was identified as lexical variation or L.V., and consisted of a count of the lexical diversity [as suggested by Chall and referred to in item (i) of Section 6.2.5.2]. For this measure each content or lexical item was counted only once. The different lexical types were then calculated as a percentage of the total number of lexical items or tokens (already counted in L.D.). This has been described by Linnarud (1975) as the Type : Token ratio. It was with this measure that the differences between the long and short texts became evident. With the long passages, the lexical variation ranged from 62% to

70%, while the shorter passages had a L.V. range from 70% to 82%.

The four very short passages (80 to 152 words) used for the reading comprehension sub-test of ELBA had an even higher L.V. range of 72% to 89%.

This seemed to suggest that when passages providing information were brief and condensed, they tended to have a higher ratio of different words. This measure of difficulty level might well be particularly pertinent to the NNS reader.

The obvious question this raised was whether all texts short or long tended to present much of the new information in the early part, or whether this very high count is specifically confined to short texts where authors of necessity must pack in the information content as densely as possible.

A fairly simple way to answer this was therefore to take the first part of the long texts and analyze them to see whether this would yield a lexical variation count that was closer to the shorter texts.

Results are presented in Table 5, and the analysis of the first part of the long texts revealed a lexical density between 41% to 56%, and a lexical variation ranging from 79% to 87%, a ratio closer to the short texts than the parent long texts of which they formed a part.

PASSAGE NO.	TITLE		LENGTH	LEXICAL DENSITY	LEXICAL VARIATION
1	CIVILISATION		LONG: 1096 words	51%	68%
			FIRST 299 words	53%	87%
2	WHAT IS LAW?	(A)	LONG: 1063 words	45%	62%
			FIRST 287 words	41%	80%
		(B)	SHORT: 287 words	44%	82%
3	URBANISATION	(A)	LONG: 1051 words	49%	70%
			FIRST 306 words	50%	81%
		(B)	SHORT: 306 words	50%	78%
4	THE LEATHERBACK TURTLE	(A)	LONG: 1203 words	51%	67%
			FIRST 323 words	56%	79%
		(B)	SHORT: 320 words	56%	76%
5	THE SCIENTIFIC STUDY OF SOCIAL DEVELOPMENT	(A)	LONG: 1071 words	49%	70%
			FIRST 328 words	49%	79%
		(B)	SHORT: 333 words	53%	70%

TABLE 5 DIFFICULTY LEVEL : LEXICAL COUNT (DENSITY AND VARIATIC

Notes:
1. Lexical word defined as nouns, verbs, adjectives and adverbs of the -ly type.
2. Lexical density = percent of lexical words to total orthographic count.
3. Lexical variation = percent of different lexical types to total lexical count.

This appears to confirm what Linnarud suggested (1975:19) 'the longer the essay, the more difficult it is for the writer to avoid repeating himself'. This was made in connection with NNS' writing, but it does indicate that it might well be true of other writing as well.

We give below the L.D. and L.V. of the ELBA reading comprehension sub-test for comparison.

PASSAGE NUMBER	LENGTH	LEXICAL DENSITY	LEXICAL VARIATION
1	80 words	45%	89%
2	104 words	50%	81%
3	152 words	43%	82%
4	99 words	51%	72%

TABLE 6 : ELBA READING SUB-TEST : LEXICAL COUNT

To briefly review this section, we used three different measures of difficulty level in order to include in this analysis the various factors Chall had suggested ought to be taken into account, and for which objective counts were possible (See Section 6.2.5.2.)

In order to compare the three measures, we used the combined ranks reported under Section 6.2.5.2 (1) as representative of readability formulae. For the deletion count, the rankings reported under Section

6.2.5.2 (2) were used; and the lexical variation count reported in Table 5 was used to rank the long and short texts since lexical density had failed to differentiate between them. The three measures are reported in Table 7 and our interest lay in seeing to what extent these measures successfully predicted the difficulty levels as they were realised in performance, and the degree of correlation obtained with the result of the experiments.

READABILITY FORMULAE	DELETION COUNT	LEXICAL VARIATION	RANKING
2A	1	2A	EASY
1	2A	4A	
4A	4A	1	TO
3A & 5A	5A	3A & 5A	
	3A		DIFFICULT
2B	2B	5B	EASY
4B	3B	4B	TO
5B	5B	3B	
3B	4B	2B	DIFFICULT

TABLE 7 : DIFFICULTY LEVEL

As is evident there was a much higher degree of agreement between the three measures for the longer texts than for the shorter ones.

These then were the various factors we took into account in our analysis of the tasks that were set to test the various hypotheses presented in this chapter.

6.3 DESIGN OF THE EXPERIMENTS

Since most of the details involved in the planning and design have already been discussed under the task variables, we will be presenting this section fairly briefly.

There were four planned experiments. Since the ordering of the pre-cloze and the multiple-choice test was important, the first experiment was primarily concerned with observing the effects of ordering.

6.3.1 Experiment I : EFFECTS OF ORDERING

Subjects were divided into two groups so that the effect of pre-reading cloze on the multiple-choice test could be investigated.

I (A)		
Group 1	Pre-reading cloze Passage 1 : 45 minutes	Text of Passage 1 with Multiple-choice questions: 45 minutes
Group 2	Text of Passage 1 with multiple-choice questions : 45 minutes	-

Because it was thought advisable to investigate the effects on cloze when it was preceded by a reading of the text, a second sample of overseas students in Egypt were assigned in random order to two groups where the cloze and multiple-choice tests were presented in counter-balanced order.

I (B)		
Group 1	Pre-reading cloze Passage 1 : 45 minutes	Text of Passage 1 with multiple-choice questions: 45 minutes
Group 2	Text of Passage 1 with multiple-choice questions : 45 minutes	Post-reading cloze Passage 1 : 45 minutes

The pre-reading and post-reading cloze tests were identical so that pre- and post- referred only to the ordering and not the construction. These two samples permitted a comparison of the effects of ordering.

6.3.2 Experiment II : PROCESS VERSUS PRODUCT, CLOZE VERSUS SUMMARY

This comprised the use of the summary as a measurement device of the product stage, and permitted a comparison with the process stage via the cloze test. The summary was a recall task requiring the subjects to write the gist of the passage. Two sample groups, both small, were used in this experiment.

Group 1	Pre-reading cloze Passage 1 : 45 minutes	Summary by recall of Passage 1 : 15 minutes

There was only a five minute interval between the reading and the recall.

Group 2	Pre-reading cloze Passage 1 : 45 minutes	Reading Passage 3A : 15 minutes	Summary by recall of Passage 3A : 15 minutes

The double bars indicate an interval of more than 24 hours. But as for Group 1 there was only a five minute interval between reading the passage and the recall of the same passage by a paraphrase of the gist.

This permitted a comparison of the process and product stages, the latter dependent mainly on working memory.

6.3.3 Experiment III : PROCESS VERSUS PRODUCT, CLOZE VERSUS MULTIPLE-CHOICE

This part comprised what we have referred to as the main experiment and the largest samples were used here. There was only one ordering.

Pre-reading cloze Passage 1 : 45 minutes	Text of Passage 1 with multiple-choice questions: 45 minutes

The purpose of the design was to enable a comparison of the two techniques reflecting the process and product of reading. A detailed analysis of cloze responses enabled us to investigate the various hypotheses proposed, as did the analysis of the multiple-choice questions based on the taxonomy.

6.3.4 Experiment IV : LENGTH AS VARIABLE

This was primarily concerned with an investigation of the effect of length on the process, and by utilising texts of two different lengths it was hoped that ample evidence would be forthcoming that would justify our insistence on the long passage as a better discriminator between varying levels of reading efficiency.

For this part the remaining eight texts were used. Each group was randomly divided into two and presented with two texts one long and one short from the same source and on the same topic by the same author. The ordering effects were cancelled out by presenting the two lengths in counter-balanced order. Four groups who participated, were each given a different pair of texts, so that the design presented below was used by all who participated in this experiment. Almost all subjects who took part in Experiment IV had also taken part in Experiment III, and in every instance Experiment III came before Experiment IV.

EXPERIMENT III	Pre-reading cloze PASSAGE A : 45 minutes	Pre-reading cloze PASSAGE B : 15 minutes
	Pre-reading cloze PASSAGE B : 15 minutes	Pre-reading cloze PASSAGE A : 45 minutes

There was no interval at all between the two parts, and apart from instructions by the supervisor to move to the second part of the experiment, there was no interruption. Both texts were stapled in the order required by the experiment, which took 1 hour.

This experiment provided more evidence of the process and permitted overall comparison between lengths of text.

6.4 SUMMARY

In this chapter, the hypotheses developed were based on the reading model. The areas proposed for experimental investigation were concerned with both the process and end-product stages, and three techniques suited to the tasks were discussed along with other task variables.

In addition because of our interest in the long text which we postulated was a more reliable discriminator between levels of reading efficiency we set up one experiment to investigate the process of reading when two different lengths of text were used.

The three techniques were discussed along with the scoring procedures adopted. In addition the selection of reading materials used for the experiments were analyzed in terms of reading difficulty by various measures that could be predictively applied.

We concluded by including the design of the four experiments that were planned and carried out, and which will be reported in Chapters Seven and Eight.

SECTION FOUR

RESULTS

Chapter Seven. Procedure and Results : Experiments I and II

Chapter Eight. Procedure and Results : Experiments III and IV

CHAPTER SEVEN

PROCEDURE AND RESULTS : EXPERIMENTS I AND II

7. GENERAL REMARKS

As we saw in Chapter Six, the techniques selected for investigating and measuring reading efficiency were chosen because they reflected the hypothetical construct of the cognitive processes postulated in the reading model.

In this chapter it is proposed to discuss two informal experiments that were carried out in addition to the four planned experiments already discussed in Section 6.3. We will also present the procedure and results of the first two experiments, ie. I and II, reserving the description and outcome of Experiments III and IV for the next Chapter.

For all experiments a 5% level of significance was accepted as providing confirmation of the hypotheses being tested.

7.1 INFORMAL EXPERIMENTS

Many of the specific hypotheses selected for testing were based on the belief that cloze is the only objective technique developed sufficiently as a measurement device that can provide us with information regarding the reading process. Since so much depended on the efficiency of the technique to reflect the influence

of not just immediate context, but more importantly, long-range constraints imposed by on-going discourse, we carried out an informal experiment to test this. The second experiment was concerned with investigating the concept of redundancy which, as we suggested, enabled the cognitive processing system to reduce input for storage to an efficient minimum. Since English exhibits this redundancy in various forms and at varying levels of linguistic analysis, we assumed that a good test of discrimination between subjects with different degrees of language proficiency would be to see the extent to which they could recognize redundancy and comprehend information from abbreviated texts, that is, reduced redundancy would not interfere with the comprehension of proficient readers.

7.1.1 Influence of Long Range Constraints on Cloze Items

In view of the doubts raised by Carroll, Carton and Wilds (1959) and MacGinitie (1961) which were referred to earlier in our discussion of the cloze technique (Section 6.2.3.1) we carried out this experiment specifically to examine the extent to which long range constraints affected the choice of words in the completion of a cloze task. MacGinitie (1961:128) had argued that 'although various sources of constraint (e.g. topic, syntax, semantic reasonableness etc.) can be distinguished conceptually, it [was] difficult in completion experiments

to separate their effects' so that a context of more than five words had relatively little effect.

Carroll et al (1959) had used twenty native speakers in their experiment to test the influence of paragraph length cues on items. They presented subjects with 20 discrete items, one a page, in random order. Four words before and five words after the deleted item provided the context. Although it was obvious that there was a connection between the items, the instructions made no mention of this. The items were chosen from three passages and the results of the twenty subjects were compared with a control group of eleven who completed the items as part of the paragraph unit. Total results were significantly different at ·001 level, although individual items were not compared. The authors concluded that 'content-bearing words' such as nouns and verbs were more likely to be affected by paragraph cues.

We adopted a similar procedure, selecting 20 items from the total of 100 that made up the cloze test based on Passage 1. These represented 4 functional and 16 content items. Within the two categories, the selection of items was random. The ratio was however deliberate since our concern was with items affected by context and usually reflected in the content items that carried much of the semantic load.

These 20 cloze items were presented on separate

slips of paper, each with five words before and five words after an underlined space representing the deleted item. The items were presented in random order to eight native speakers who acted as subjects. Each of the eight subjects was presented individually with the task, and was orally instructed to complete the 20 frames by supplying the missing items.

If contextual cues were of little assistance beyond the five word range, native speakers should experience little or no difficulty in producing an 'acceptable' response for the missing item. If however, their performance indicated that items were difficult to complete accurately, this would be taken to suggest that context was of considerable help, if not necessary for the task.

Of interest was the reaction of the eight NS to the task. Whereas six of them accepted the instructions at face value, producing responses that were unacceptable in terms of the original context, two of the NS were observed after the first couple of frames, to turn back and then shuffle through the slips of paper. Their results differed from the other six, and indicated very clearly their efforts to impose a common frame of reference for all the items. Their scores were correspondingly higher indicating that proficient readers can impose meaning even when provided with only fragmentary evidence of the original text. The task as presented is contained in Appendix 1.C.

Results are presented below, and for comparison, scores of eight non-native speakers who completed the items with benefit of context. These 8 NNS formed part of the group which came in the Upper Third category of the cloze test from which the items were taken. Their results for these items were extracted from their total scores obtained on the passage as a whole.

Both individual and item scores are presented below.

SUBJECT NO:	1	2	3	4	5	6	7	8	TOTAL SCORE
8 NATIVE SPEAKERS	12	10	4	3	1	5	3	3	= 41
8 NON-NATIVE SPEAKERS	11	14	17	17	13	14	15	13	= 114

TABLE 8 RESULTS OF CLOZE SCORES: NS VERSUS NNS

Mean for the native speakers : 5·125

Mean for the non-native speakers : 14·25

A t-test for independent samples confirmed that the difference between means was significant at above the ·05 level. (Obtained value t = 5·935, and value needed for significance at ·05 level, t = 2·145)

We give below scores obtained on individual items by the two groups and as a comparative statistic, the facility value obtained by the entire Upper Third

sample of 63 non-native speakers who sat the cloze test based on this passage. These figures are drawn from details of their results contained in Appendix 4.C.1. for Passage 1. Item numbers are given as in the original passage.

FUNCTIONAL ITEMS				
ITEM NO:	11	14	51	56
NS SCORES	2	2	4	2
NNS SCORES	7	8	6	8
FACILITY VALUE:	·70	·87	·67	·95

CONTENT ITEMS (10 nouns + 6 adjectives)																
ITEM NO:	2	12	21	23	24	37	44	54	61	66	72	73	83	84	89	97
NS SCORES	2	4	2	2	2	3	1	2	2	3	2	0	1	1	4	0
NNS SCORES	1	5	4	5	6	7	6	6	6	5	3	8	8	5	7	3
FACILITY VALUE	·38	·49	·44	·43	·49	·75	·79	·57	·41	·65	·59	·71	·89	·75	·81	·35

TABLE 9 **RESULTS ON 20 ITEMS : NS VERSUS NNS**

The results when calculated for both functional and content categories were fairly conclusive, being significantly different at above ·05 level for the two groups.

Native speakers scored 31·25% of functional, and

24.125% of content items. NNS scored 90.625% of functional, and 66.41% of content items, confirming the assumption that long-range constraints did help in the successful completion of deleted items as proved by the higher scores of non-native speakers. Items were classified according to use in the specific contexts in which they occurred and were scored by the acceptable word method used for the cloze test. Although non-native speakers performed the cloze test under timed conditions, native speakers had no such limits imposed. This informal experiment showed very clearly that context was crucial in determining the distributional choice of lexical items so that cloze tasks did utilise both short and long-range constraints when determining selection.

7.1.2 Influence of Redundancy

The second informal experiment was aimed at examining the concept of redundancy. In the formal experiments, all deletions were indicated by a standard underlined space. We assume that in cloze most, if not all, items deleted are recoverable from the remaining text because of the in-built redundancy present in the language. We were interested in approaching the task from another angle to investigate the ease with which abbreviated texts could be expanded to match the original. We judged that sensitivity to language form and meaning would be evidenced in subjects' attempts to recompose the text.

We adopted a strategy used by Tuinman and Gray (1972) where the deleted words were not replaced by an underlined space or other marking device. The text was abbreviated by 10% by leaving out the words deleted in the cloze test based on Passage 1. Two non-sequential paragraphs were selected from the passage and the running text presented with the following written instructions. The Text is contained in Appendix 1.D.

Instructions : This is an abbreviated text.
(1) Reconstruct the original text and (2) indicate where, according to you, the missing item(s) were initially included.

Sixteen subjects, eight native speakers - NS - and eight NNS took part in this experiment. None of them had taken part in either the earlier or more formal experiments to be reported later. Once again 20 items were deleted from two paragraphs of Cloze Test Passage 1. The two paragraphs were not consecutive, one coming from the early part and the second comprising the last part of the original text.

The results are reported below. Although the place where the text was missing was indicated, unfortunately the items were not inserted by all the subjects' so that this part of the experiment could not be analysed. There was a marginal difference between the two groups with native speakers averaging a slightly higher mean. Each indication of a missing item, if

correct, was scored, so that the total possible was 10 marks a paragraph. The NNS were all post-graduate students at Edinburgh University which might explain the very small difference between the groups. The items not marked as missing from the text were adjectives and a noun that formed part of a list [item 3].

No time limit of any sort was imposed for the task which was unsupervised.

ITEM(Form class)		NON-NATIVE SPEAKERS								TOTAL	NATIVE SPEAKERS								TOTAL
		1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8	
1	determiner	✓	✓	–	✓	✓	✓	–	–		–	✓	✓	✓	✓	✓	✓	✓	
2	noun	✓	–	✓	✓	✓	–	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
3	noun	–	–	–	–	–	–	–	–		–	–	–	–	–	–	–	–	
4	determiner	–	–	–	–	✓	–	–	✓		–	–	–	–	–	–	–	–	
5	conjunction	✓	✓	✓	✓	–	–	✓	✓		✓	✓	✓	✓	–	✓	✓	✓	
6	noun	✓	–	✓	✓	✓	–	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
7	preposition	–	–	✓	–	✓	–	✓	✓		✓	✓	✓	✓	–	–	✓	✓	
8	preposition	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	–	✓	
9	noun	✓	✓	✓	✓	✓	–	–	✓		–	✓	✓	✓	–	✓	✓	✓	
10	auxiliary	✓	–	–	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
	TOTAL	7	4	6	7	8	3	6	8	= 49	6	8	8	8	5	7	7	8	= 57
11	noun	✓	✓	✓	–	✓	–	✓	–		✓	✓	✓	✓	–	–	–	✓	
12	pronoun	✓	✓	✓	✓	✓	–	✓	✓		✓	✓	✓	✓	✓	✓	✓	–	
13	conjunction	✓	–	✓	✓	✓	✓	✓	✓		–	✓	✓	–	✓	✓	✓	✓	
14	preposition	✓	–	✓	✓	✓	–	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
15	preposition	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
16	conjunction	✓	✓	–	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
17	verb	✓	–	✓	✓	–	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
18	adjective	–	✓	–	–	–	–	–	–		–	–	–	–	–	–	–	–	
19	noun	✓	–	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
20	adjective	–	–	–	–	–	–	–	–		–	–	–	–	–	–	–	–	
	TOTAL	8	5	7	7	7	5	8	7	= 54	7	8	8	7	7	7	7	7	= 58
INDIVIDUAL TOTALS		15	9	13	14	15	8	14	15		13	16	16	15	12	14	14	15	

TABLE 10 RESULTS OF RECONSTRUCTION OF ABBREVIATED TEXT : NNS VERSUS NS

The mean for non-native speakers was 12.875 versus native-speakers' mean of 14.375. (Total possible : 20 marks)

Results showed that both groups were sensitive to language. It is possible that further experimentation with this type of test might prove rewarding since it would force the subject to judge existing text for deviancy. And by requiring reconstruction and recomposition it would reveal evidence of both grammatical and semantic competence. This was our intention but as reported, the second part of the exercise was overlooked by most subjects.

What is of special interest is that apart from the correct location of deletions, there were many more items indicated as missing from the text by NNS as compared to NS.

For Paragraph 1 : 5 additional deletions (NNS) versus 1 (NS),

For Paragraph 2 : 6 additional deletions (NNS) versus 3 (NS).

The results of these two informal experiments confirmed our view that cloze was not just a test of the influence of immediate linguistic elements. When a sufficient number of items were deleted, cloze provided a valid measure of the way in which comprehension of earlier parts of discourse could and did influence understanding since long-range constraints were used to help complete the deleted items. The ability to recognize and utilise redundancy was essential for efficient reading, and a test that evaluated this skill was evidently of use.

7.2 FORMAL EXPERIMENTS

We come now to a description of the more formally planned experiments.

7.2.1 Experiment I : Effects of Ordering

Two experiments were conducted to assess the effects of ordering of tests as a possible source of variation. Because we were interested in the initial interaction between reader and text at first exposure, we had decided that pre-reading cloze would be the first test when more than one technique was used based on the same material. This meant that multiple-choice questions on the same text would form the second part of the experiment. From our reading model, and reflecting the hypotheses developed we assumed a null hypothesis regarding the influence of cloze on multiple-choice. Since different cognitive stages were being assessed i.e. one the process and the other, the utilisation of the product of reading, we hypothesised that ordering would not significantly affect the results. This led to the first try-out. As stated earlier a 5% level of significance was adopted.

7.2.1.1 Experiment I (A)

This first experiment was conducted with two groups drawn from the non-native speaker population attending courses in their first year at Edinburgh University.

GROUP A was required to do two tests:

(i) Pre-reading cloze based on Passage 1, followed by

(ii) A reading of the same text unmutilated together with a list of 20 multiple-choice questions.

Subjects in this group were informed that they would be required for 1.1/2 hours. A short coffee break was given between the two tests permitting collection of Test 1 papers, and distribution of Test 2 without loss of testing time.

Instructions for the cloze test were read out orally, and printed on the test booklet. The oral instructions were as follows : [See Appendix 1.A.3. for complete details].

SECTION 1 Time : 45 minutes

'This is an exercise in reading comprehension. It consists of a long passage with blank spaces. You are required to fill in the blank spaces using only one word in each space. You will have to guess what the missing word is likely to be, taking into account the language and meaning of the passage. Work as quickly as you can. If you are not sure and cannot make a guess, move on to the next blank space. Don't waste time. You may be able to fill in the missing word later when you have read more of the passage.

If you do not understand what you have to do, please ask now. You will not be allowed to ask any questions once the exercise has started.

Now write your name and date at the top. You have 45 minutes for this exercise.

Are there any questions?'

(Pause. Please repeat or clarify if necessary. If there are no questions,) 'Please begin now'.

On each of the test booklets, the following instructions were printed : [See Page 9, Appendix 1.B.1.1]

> Complete the following passage by filling in the blank spaces as you think best. Only <u>one</u> word per blank space should be used. You have 45 minutes for this exercise.

At the end of 45 minutes the cloze test was collected. Similarly there were instructions both oral and written, for the multiple-choice test [See Appendix 1.A.4. and 1.B.1.6.]

<u>GROUP B</u> did only the multiple-choice test comprising the same unmutilated text and 20 questions used for Group A. In addition, an information sheet was also supplied asking for details regarding their first language, and the number of years they had studied English. It was expected that the latter data might provide useful information. Unfortunately there was so much variation reported regarding the type of English instruction received, that this part of the test was eventually discarded. Some of the subjects had received formal instruction through English, others about English, some had received private tuition, others had attended language classes. Some had studied one or two subjects in English, others had changed educational institutions so that there was no clear answer to the question asked. A copy of this Evaluation Form is given in Appendix 1.A.6.

To confirm that both groups were drawn from the same population, an F-test, the variance-ratio test was performed.

Obtained value F = 1·2. Value needed for 5% level of significance, F = 2·31 for a two tailed test.

Having confirmed the homogeneity of the population from which the two samples were drawn, we performed a t-test for independent samples on the results of the multiple-choice questions. As the results in Table 12 indicate, the observed difference in means could have occurred by chance which enabled us to accept the null hypothesis that multiple-choice questions were not significantly affected by being preceded by a cloze test on the same passage.

The mean, standard deviation, and standard error of the mean are reported for all results.

TEST	MEAN	S.D.	S.M.	CORRELATION COEFFICIENT	
				CLOZE (Acceptable)	MULTIPLE-CHOICE
CLOZE (Exact)	50.2	10.72	2.14	·91*	·75*
CLOZE(Acceptable)	69·48	14	2·81	-	·82*
MULTIPLE-CHOICE	14·44	3·27	·65	-	-

TABLE 11 GROUP A RESULTS N = 25 (* p < ·05)

The results of GROUP B which did just the multiple-choice test are given below along with those of GROUP A on the same test.

	MEAN	S.D.	S.M.	
GROUP B : N = 24	12·71	2.97	·61	F = 1·2
GROUP A : N = 25	14·44	3·27	·65	t_1 = 1·9,[1] p > ·05

TABLE 12 GROUPS A & B : RESULTS OF MULTIPLE-CHOICE TEST

Since the null hypothesis was retained, we combined the two groups for purposes of analysis of the multiple-choice test. The results for the combined sample are reported here :

N = 49 Mean = 13·59, S.D. = 3·21, S.M. = ·46

Details regarding individual results are contained in Appendix 4.A.1. Item analysis results will be discussed later along with other data on the multiple-choice versions. Suffice it to say here that applying the Kuder-Richardson (20) formula using the proportion of passes and failures for each item with 49 subjects, the coefficient of reliability obtained was r = ·71.

Validity of the multiple-choice test was obtained by correlating results of 25 subjects who had also sat the ELBA (Ingram et al), our external criterion measure. Since detailed results of subjects were available, a correlation was made both with the Reading Comprehension sub-test and the total scores. As reported earlier, the reading comprehension sub-test comprised 20 multiple-choice questions based on four passages. (See Section 3.3.5. for more information)

---

1. Value needed for significance at ·05 level with 47 df was t = 2·018, Lewis (1972:112)

N = 25	ELBA (TOTAL)	ELBA(Reading Comprehension)
MULTIPLE-CHOICE	.66 *	.71 *

TABLE 13 : CORRELATION COEFFICIENT OF MULTIPLE-CHOICE 1 & ELBA

This experiment confirmed our hypothesis regarding the negligible effects of ordering on the multiple-choice test as a result of being preceded by a cloze test. It also confirmed that the multiple-choice test was a reliable and valid measure of reading comprehension.

7.2.1.2 Experiment I (B)

The second experiment inevitably was concerned with effects of ordering on cloze. Sixteen Egyptian teachers of English enrolled at a Masters' course in Alexandria, Egypt took part in this experiment doing the same two tests reported in 7.2.1.1. The multiple-choice test however had 25 questions, five more than used earlier. In addition items were revised as a result of item analysis performed on the earlier version, but details of the tests will be reported later. The subjects were assigned in random order, doing both tests in counter-balanced order to the following design :

GROUP A : N = 8 (i) Pre-reading cloze based on Passage 1
followed by (ii) Multiple-choice questions on Passage 1

GROUP B : N = 8 (i) Multiple-choice questions on Passage 1
followed by (ii) Cloze test based on Passage 1

For both groups, the tests were identical. Once again we adopted the null hypothesis and a 5% level of significance for acceptance or rejection.

The two tests were however a week apart and this time subjects were not told they would be tested again. Because of the time interval between the tests, instructors were specifically advised to say nothing that would encourage subjects to exchange information or rehearse content. Both groups sat the test in the same room, and both sets of instructions were read out on both occasions. Individual results are recorded in Appendix 4.A.2., but group results are presented below in Table 14.

	GROUP A (N=8)			GROUP B (N=8)		
TEST	MEAN	S.D.	SM	MEAN	S.D.	SM
CLOZE (Exact)	43·63	7·56	2·67	50·13	8·58	3·03
CLOZE(Acceptable)	61·63	9·38	3·32	67·38	11·89	4·2
MULTIPLE-CHOICE	16.38	2	.71	13.13	3.48	1.23

TABLE 14 : RESULTS OF EXPERIMENT I (B)

A t-test for independent samples was carried out using the means and our hypothesis was confirmed that cloze scores were not significantly affected by the ordering.

Obtained values for t : (i) With exact cloze, t = 1·61
(ii) With acceptable cloze, t = 1·07
(iii) With multiple-choice , t = 2·29

Value needed for significance at 5% level with 14 df, t = 2·145. [As recommended for a sample of less than thirty, (n - 2) degrees of freedom were used, Lewis (1972:175)].

Results for this multiple-choice test were however significant at 5% level. We have no idea whether discussion

among subjects occurred between the tests outside the classroom. Combined results for the cloze are reported in Table 15 along with reliability coefficient. The only external criterion available for this sample was a test administered by the Institute some three months prior to the experiment. The test comprised three parts:

(i) Listening comprehension - 20 marks

(ii) Aural perception (phonemic discrimination) - 10 marks

(iii) Continuous writing - 70 marks, Total = 100 marks.

N = 16	MEAN	S.D.	SM	RELIABILITY	CORRELATION	
					CLOZE (Accep)	INST.T
CLOZE (Exact)	46.88	8.5	2.13	.74	.98*	.3
CLOZE (Acceptable)	64.5	10.76	2.69	.84	-	.29

TABLE 15 : RESULTS OF CLOZE TEST : EXPERIMENT I (B)

As we can see correlations with the Institute test were low. We used candidates' total scores on the internal test for this. Although Experiments I (A) and (B) had used two samples widely separated by location, we were fairly certain that both represented the overall population we were interested in for this thesis. We therefore thought it worth while comparing the results of their performance on the cloze test since this was sat by both groups.

7.2.1.3 Comparison of Samples

The variance-ratio (F test) between the two samples confirmed the homogeneity of variance assumption. The t-test for independent samples using means confirmed that differences could have occurred by chance in more than 5% of cases.

		EXPERIMENT I (A) N = 25	EXPERIMENT I (B) N = 16	
CLOZE (Exact)	MEAN	50·2	46·88	F = 1·59[1]
	S.D.	10·72	8·5	t = 1·05
	SM	2·14	2·13	p > ·05
CLOZE (Acceptable)	MEAN	69·48	64·5	F = 1·69
	S.D.	14	10·76	t = 1·21
	SM	2·81	2·69	p > ·05

TABLE 16 : COMPARISON OF CLOZE : EDINBURGH VERSUS EGYPTIAN SAMPLE

The smaller standard deviation for the Egyptian sample was understandable because of their greater conformity in level of language ability since all subjects had passed the same entrance test and were classmates on the same course. Their acceptance on the course was based on test results unlike the Edinburgh group where all of the subjects were admitted despite variations in the results of their entrance test, ELBA.

What both experiments confirmed was the ability of the cloze test i.e. Cloze Test Passage 1, to discriminate between levels of reading ability both in Britain and overseas provided the same population was used, that is, non-native speakers undergoing academic courses with instruction in English. We turn now to our second experiment.

7.2.2 Experiment II : Process versus Product, Cloze versus Summary

This was an experiment designed to measure the

1. A value of F = 2.7 with (24,15) df was needed for significance at .05 level, as was a value of t = 2.023 with 39 df.

process and product stages of reading using two techniques that we had postulated were valid measures of the different stages. The product in this experiment was a far more direct measure than in Experiment I. It was hypothesised that information content was organized and stored in working memory, and if subjects were to produce evidence of reading comprehension this was the area to be explored. A technique requiring subjects to re-encode contents of working memory would elicit evidence of their understanding always tempered by ability to re-encode meaning into the foreign language.

Our technique required recall; but since a paraphrase could be used, the demands on memory which verbatim recall would have imposed, were avoided. This open-ended task had all the attendant difficulties associated with its use as a measurement device. To introduce a measure of reliability the texts were analyzed and a list of 25 ideas drawn up prior to correction. As reported earlier we had initially planned to use summaries written by native speakers but these failed to produce any clearly defined criterion that we could apply to NNS. To obtain the list of ideas, the text was analysed using a fairly crude scheme that categorized all sentences in terms of discourse function. Some sentences made assertions, others elaborated the assertion, some produced evidence of the facts asserted, while some functioned as links between ideas or parts of the text contributing more to coherence than development of information. The summary was abstracted by using the information content of the

'assertions' aided by 'evidence'.

Details of the system, along with definitions of terms used are provided in Appendix 3. Also given is the analysis of texts used in this experiment and the list of ideas that were drawn up for purposes of scoring the summary. Two groups were used in this experiment. All subjects also did the cloze test based on Passage 1, before the summary question.

7.2.2.1 Group A : First Sample

The first sample was small and comprised eleven students who had just matriculated at Edinburgh University. They were however in need of remedial tuition and were receiving instruction in English. They were given the following instructions which were read out.

> 'There are two parts to this test. Part I will be done first and the papers collected after 45 minutes. You will then do Part II. In this test you will be required to write a summary of the passage you use in Part I. Detailed instructions for each part will be given separately'.

For the pre-reading cloze test, the same instructions already reported in Section 7.2.1.1. were read out. At the end of 45 minutes, the test papers were collected. Single sheets of blank paper were then distributed and the following instructions read out aloud.

> Part II: 'Now write your name at the top of the sheet. You have 15 minutes to write a summary of the passage you have just read. The summary may be in note form with the main points only. You needn't use complete sentences'.

This whole experiment for this sample lasted just over an hour. Individual scores are given in Appendix 4.B.1. but group results are presented below. Since all eleven subjects had sat the ELBA test a few days previously we used this and the ELBA Reading Comprehension (R.C.) scores as well for purposes of correlation.

Summary (N = 11) : Mean = 7·73, S.D. = 2·65, SM = ·8 Reliability was ·34, fairly low in view of the subjective marking.[1] Correlations are given below with the criterion measure, and the cloze test based on the same text.

N = 11	CLOZE (Exact)	CLOZE (Acceptable)	ELBA (Total)	ELBA (Reading Comp).
SUMMARY	·65*	·75*	·06	·66*

TABLE 17 : CORRELATION COEFFICIENT OF SUMMARY & ELBA: GROUP A

Even with this very small sample it was still evident that when the two measures were obtained using the same text, there was a significant correlation. The contents of working memory were obviously affected by how the input was processed via cloze. What is unexpected is the very low correlation with the total ELBA scores, the latter comprising oral comprehension, vocabulary, grammar and reading comprehension scores. Correlation with ELBA (R.C.) Reading Comprehension sub-test was satisfactory.

---

1. Downie and Heath, 2nd edition (1965:222): Saupe's formula.

7.2.2.2 Group B : Other Sample

Two other samples were used, one of thirteen drawn from NNS at Edinburgh University who like the previous group had just matriculated but were attending remedial classes in English. The second sample of fifteen subjects were all admitted to various courses at British universities but were at the time of the experiment undergoing full-time instruction at a language school in Haddington near Edinburgh. Both samples did the same tests in the same order. They were first asked to do the Pre-reading cloze test based on Passage 1 and given exactly the same instructions as others doing the test in earlier experiments. Two days later in both samples, subjects were given the second part of the experiment which comprised a reading unit to be followed by summary-writing.

The following instructions were read aloud. (See Appendix 1.A.1.)

'You have 15 minutes to read this passage, after which it will be taken back. You will then be required to write a brief summary of the passage'.

Passage 3A was chosen for this experiment. After 15 minutes the passage was collected and single sheets of blank paper distributed as in the previous experiment just reported. The same instructions were then read aloud as for Group A Part II. (See Section 7.2.2.1)

This experiment lasted just over half an hour. In order to combine results of the two samples an F test was performed confirming that they were drawn from the same population.

Value needed for significance F = 3·23 (with 14,12 df).
Value obtained from exact cloze : F = 1·73, from acceptable cloze : F = 2·14. (with 14, 12 df)
Value needed for significance F = 3·06 (with 12, 14 df)
Value obtained from summary tests : F = 2·01 (with 12, 14 df)

The combined results for twenty-eight subjects are presented below, individual scores being reported in Appendix 4.B.2. Since the thirteen Edinburgh subjects had sat the ELBA, and fourteen of the Haddington sample the EPTB, these results were used for external validity measures.

Reliability for the summary test is difficult to assess. No attempt was made to obtain a second marker and it is accepted that this was subjectively scored although the system adopted did avoid purely impressionistic marking. Reliability here was higher than for Group A because of the larger sample used, and results are presented in greater detail below.

	N = 28 CLOZE (Exact)	N = 28 CLOZE (Accept)	N = 13 ELBA (Total)	ELBA (R.C.)	N = 14 EPTB (Total)
SUMMARY P.3A	·58*	·64*	- ·19	·51	·53*
CLOZE(Exact)	-	·97*	·32	·62*	·76*
CLOZE(Accept)	-	-	·29	·78*	·79*
N = 28	MEAN	S.D.	SM	RELIABILITY	
SUMMARY P.3A	10·57	3·55	·67	·65	
CLOZE(Exact)	28·5	9·13	1·73	·78	
CLOZE(Accept)	37·18	13·29	2·51	·90	

TABLE 18 : RESULTS OF EXPERIMENT II : GROUP B

What is interesting is the absence of any kind of correlation between the summary results and total performance at the proficiency test ELBA in this second sample as well. (Also see Table 17). With the Reading Comprehension sub-test, results almost approached significance, a value of ·53 needed at ·05 level. The second test of proficiency, the EPTB, did however reveal a significant degree of correlation with the summary. Correlations between the two techniques, that is, cloze and summary-writing were higher when based on the same text as in 7.2.2.1 than when based on different texts as in this second sample, but details will be discussed later in Chapter Nine.

Cloze results when scored by the acceptable word method which reflects comprehension unrestricted by the need for exact word used by the author, correlated higher with summary for both samples.

7.3 SUMMARY

To recapitulate, in this Chapter we have presented results of two experiments conducted informally with small groups. They were aimed at investigating the effects mainly of long-range contexts on cloze items, and the effects of reducing redundancy by omitting overt markers to warn readers of the presence of deleted text. In the first experiment the significantly poor performance of native speakers as compared to non-native speakers can only be interpreted as positive evidence of the enormous advantage

provided by increased context to NNS. The ease with which NS and our proficient NNS estimated the presence of missing items in the second experiment, proved very clearly the sensitivity of both groups to language in use, and the presence of redundancy which enabled them to comprehend and follow discourse despite a reduction of ten percent of the text. Of particular interest was the discovery that NNS seemed to feel the need for the inclusion of more items than were originally deleted. Although not explored further this indicated interesting avenues for further experimentation. The two more formally conducted experiments laid the groundwork for the main experiment to be described in the next Chapter.

The first of these two studies was designed to investigate the possible influence of ordering on performance in two tests based on the same material. Results confirmed the hypothesis that order effects were not significant when the techniques used in the tests activated different stages of the reading activity associated with process and product. This enabled us to ignore serial order of tests as a possible variable affecting the results.

The second study also investigated the process and product stages of reading by a comparison of cloze and summary-writing based on the same and on different texts. And although results produced evidence that appeared to confirm the definability of the two stages, the unreliability of the summary task as a measuring

instrument served to lower the validity of any inference we might care to make on the basis of these results. This served to underline the necessity of using devices that could be proven to have statistical reliability and validity, and hence the necessity for further experimentation using different techniques.

CHAPTER EIGHT

PROCEDURE AND RESULTS : EXPERIMENTS III AND IV

8. GENERAL REMARKS

As the hypotheses postulated in Chapter Six make clear one of the major aims of this investigation was to obtain evidence of the process and the product of reading. By utilising a common text, the input variable was controlled so that the output would be primarily a result of different techniques that measured different aspects of the cognitive processes involved in reading. Experiment III was planned for this purpose, whereas in Experiment IV, length as the input variable was investigated by using only one technique to measure the processing of English during reading at first exposure to the text.

8.1 EXPERIMENT III : THE PROCESS AND THE PRODUCT OF READING

This comprised what we have referred to as the 'main' experiment. As our design indicated the purpose was to obtain reliable and valid evidence of :

(i) comprehension of the text at first exposure while it was in the process of being read, measured via a cloze test.

(ii) the exercise of various skills and strategies to obtain answers to specific questions aimed at evaluating the comprehension of the text after reading had been completed. This was via a set

of multiple-choice questions framed in accordance with a taxonomy of abilities non-native speakers were expected to need for successful reading comprehension at this advanced level. This tested the product of reading.

(iii) the extent to which the two measures correlated with each other.

Each of these three will be dealt with separately.

8.1.1 The Cloze Test : Evaluation of the Reading Process

As reported earlier the test was based on Passage 1, a 1096-word passage on the subject of the growth of civilisation. The cloze test used the 'any' word method deleting every tenth word. The first and last sentences were however left unmutilated, and each deleted item was represented by a standard underlined length of 14 spaces.

Instructions to subjects were read out orally and were printed in the test booklet. The time allowed was restricted to 45 minutes which was estimated as sufficient for the demands made, provided the students were proficient enough in the language to comprehend the text. Carroll et al (1959:83) had reported that according to their calculations, a 9-minute time limit for a 20-item word deletion cloze passage, yielded a reliable estimate. We are reminded that his subjects were non-native students at a university who sat a cloze test also of the 'any' word type with every tenth word deleted. Try-outs with both native and non-native speakers convinced us of the

reasonableness of the time limit provided, an estimate that our experiment was to confirm. Test instructions, test booklet and individual results are all contained in the Appendices. [Test materials in Appendix 1.B.1, results in Appendix 4.C.3.]

A total of 275 subjects took part in this experiment, of whom 190 were non-native speakers in Britain. They will be referred to hereafter as the 'British' sample. They were drawn from the student population of non-native speakers who had either just matriculated or, although admitted to a university, were undergoing intensive tuition in English just before university courses began. They represented students from over thirty countries, with varying educational backgrounds who were all in Britain for the first time to undertake higher level courses at an academic institution.

The remaining 85 students were drawn from three overseas centres which cooperated with this writer in running the experiment. Since an analysis of variance revealed that differences among them were significant at the ·05 level, the result of the three overseas samples will be reported independently. They comprised:

(i) 53 students from Nuremberg in West Germany

(ii) 16 students from Krefeld in West Germany

and (iii) 16 students from Alexandria in Egypt.

Regarding reliability estimates, it was first proposed to use the Kuder Richardson Formula (20) which

yields a coefficient of internal consistency, and as such, has some of the limitations of the split-half method. Carroll et al (1959) however had advised against the use of measures based solely on internal consistency as being inappropriate for continuous passages, since the two halves were not independent. It was therefore decided to use the standard error of measurement which represents 'the standard deviation of a sample of scores of an individual about his true score [if] it were possible to administer the same test repeatedly to the same individual, without any changes occurring in the individual', Downie and Heath (1965:211). The authors quote a formula by Lord (1959) which shows how the standard error of measurement can be obtained from the following formula : SE = ·432 k where 'k' is the number of items in the test. They further report that Saupe (1961) applied this formula to the KR (20) and reduced the latter to a formula for obtaining the reliability coefficient. It was this statistic that combines the standard error of measurement with the internal consistency measure based on number of test items, that we adopted for all the tests in the experiments.

$$\text{Reliability} = \frac{k}{k-1}\left[1 - \frac{\cdot 19k}{SD^2}\right]$$ Downie and Heath (1965:222)

Regarding the question of validity, construct and content validity had been established through the discussion of the reading model and in our analysis of the deleted items when evaluating the cloze technique in Section 6.2.3.1.

What this experiment enabled us to do was establish the pragmatic validity of the tests in practice, by correlation with two external measures of proficiency.

8.1.1.1 The British Sample : N = 190

It was with subjects from this sample that the correlation figures are reported with the ELBA and EPTB. In addition, 34 others of the sample had sat a test set by the language school they were attending, and results from this were also used for correlations. The test is identified as LANG.T. All data are rounded to two decimal points.

Results for the cloze were scored by exact and acceptable word methods. Implications of all the findings will be discussed in Chapter Nine.

	MEAN	S.D.	SM	RELIABILITY
CLOZE (Exact)	33.04	13.38	.97	.90
CLOZE(Acceptable)	43.92	18.76	1.36	.96
ELBA (N = 45)	64.59	18.99	2.83	.86
EPTB (N = 51)[1]	35.34	5.23	.73	-
LANG.T. (N = 34)	67.29	13.47	2.31	NOT AVAILABLE

TABLE 19 : RESULTS OF CLOZE TEST (N = 190)

Reliability estimates for EPTB and the Language Test were not available for comparison.

---

1. EPTB reports results in standard scores. Normative statistics are S.D. = 6, Mean = 40.

N = 190	CLOZE (Acceptable)	ELBA (N=45)	EPTB (N=51)	LANG.T. (N=34)
CLOZE (Exact)	.98*	.84*	.66*	.47*
CLOZE(Acceptable)		.87*	.72*	.52*

TABLE 20 : VALIDITY COEFFICIENTS FOR CLOZE TEST, PASSAGE 1

All correlations were significant at more than .05 level using the Pearson Product moment correlation coefficient formula.

Figure 11 shows the frequency distribution for this sample with a satisfactory spread of scores ranging from 9 to 85 by the acceptable word method. There is a larger spread at the upper end of the scale with a slight negative skewing showing a bunching left of centre. When the group was divided into thirds on the basis of total acceptable scores, and an F test — the variance - ratio test — run to determine whether they formed three distinct groups, the Upper Third was found to be significantly different at the .05 level. The Middle and Lower third groups presented a distribution within chance level. Details are given below.

Upper and Middle thirds: F = 2.94

Upper and Lower thirds: F = 2.45

Middle and Lower thirds: F = 1.20

Value needed for significance at .05 level, F = 1.52 for a one-tailed test.

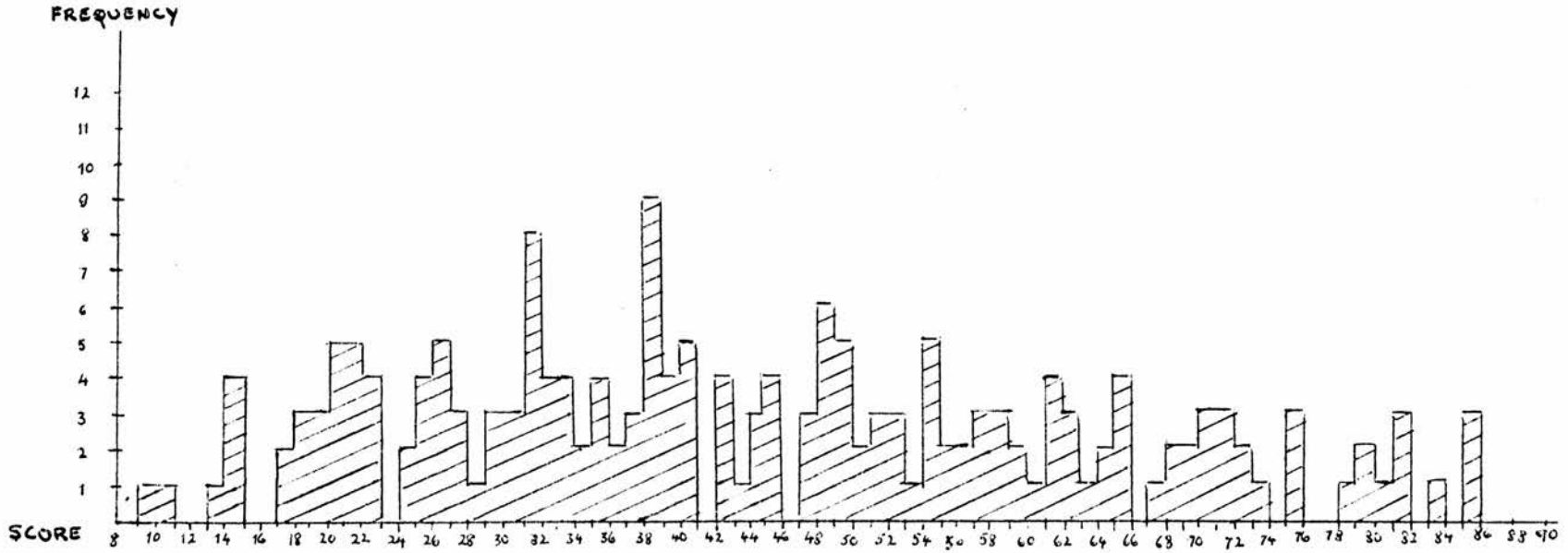


Fig. 11. FREQUENCY DISTRIBUTION OF CLOZE TEST : PASSAGE 1

BRITISH SAMPLE N = 190. Scored by acceptable word method

Since there were 100 deletions, a total of 19,000 responses were obtained from this sample. These were analyzed by exact and acceptable word methods; those that did not conform to the criteria laid down for the latter were marked incorrect. Appendix 1.B.1.2 contains the list of alternative responses that were considered acceptable. All unmarked items were scored as omitted.

Table 21 presents the details of the analysis along with correlations between the two systems of scoring.

BRITISH SAMPLE (N = 190)		
CORRECT (Acceptable):	8345,	Mean = 43·92
INCORRECT:	6580,	Mean = 34·63
OMITTED:	4075,	Mean = 21·45
TOTAL:	19000	

UPPER THIRD N = 63		
CORRECT (Exact): 3040	CORRECT(Acceptable): 4156	
r = ·86*		
Mean = 48·25	Mean = 65·97	
	INCORRECT: 1764,	Mean = 28
	OMITTED: 380,	Mean = 6·03
	TOTAL: 6300	

MIDDLE THIRD N = 64		
CORRECT (Exact): 2029	CORRECT(Acceptable): 2690	
r = ·88*		
Mean = 31·7	Mean = 42·03	
	INCORRECT: 2403,	Mean 37·55
	OMITTED: 1307,	Mean = 20.42
	TOTAL: 6400	

LOWER THIRD N = 63		
CORRECT(Exact): 1209	CORRECT(Acceptable): 1499	
r = ·95*		
Mean = 19·19	Mean = 23·79	
	INCORRECT: 2413,	Mean 38.30
	OMITTED: 2388,	Mean = 37.90
	TOTAL: 6300	

TABLE 21 : DETAILS OF CLOZE RESPONSES FOR BRITISH SAMPLE

The results of the total test based on 100 items were further divided into ten blocks based on performance of ten items per block, so that by obtaining the mean for each block of ten items, a comparison could be made between the performance of subjects in the upper, middle and lower third groups as they progressed through the passage. These mean scores for correct (by acceptable word method), incorrect and omitted items are given in Appendix 4.C.1. (page 141).

To enable a visual comparison of performance, the results of upper and lower third groups were translated into graphs for convenience to help in identifying the varying patterns of performance. These are given in Figures 12, 13 and 14 on the following page.

Of interest is the increasing variation with length between the performances of the two groups especially as scored for the correct and omitted items.

CLOZE TEST PASSAGE 1 BRITISH SAMPLE N = 190

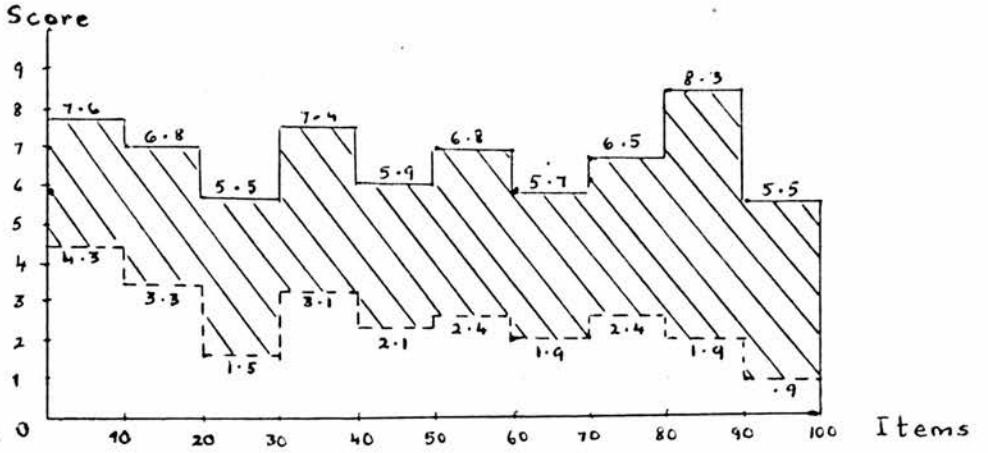


Fig. 12 CORRECT SCORE

—— Upper Third Mean = 65·97
---- Lower Third Mean = 23·79

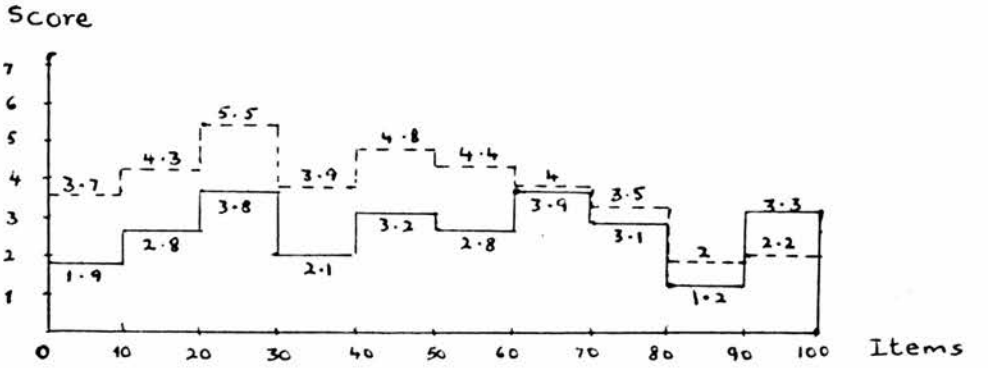


Fig. 13 INCORRECT SCORE

—— Upper Third Mean = 28
---- Lower Third Mean = 38·30

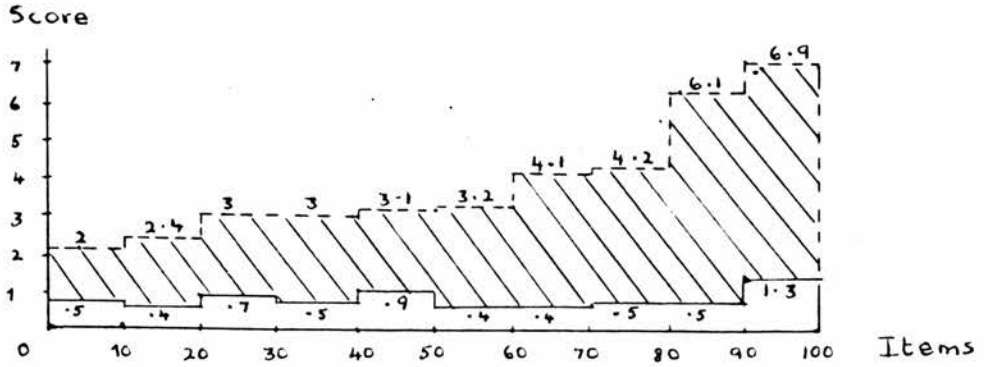


Fig. 14 OMISSION SCORE

—— Upper Third Mean = 6·03
---- Lower Third Mean = 37·90

In addition to this method, two other kinds of analyses were applied to elicit more information from the data obtained.

(i) The first was an item analysis to find out which items succeeded in discriminating between the better and less proficient readers. This was done by calculating two statistics for each item, the FV or facility value to determine ease or difficulty for the whole group, and DI or discrimination index to determine 'good' efficient items. Details are contained in Appendix 4.C.1.1., and the criteria adopted for evaluating an item as 'good' were:

FV to be between ·25 and ·85
DI to be ·25 and above

Of the 100 items in the test, 64 satisfied this criteria.

(ii) These 64 items were then classified according to word class to see which type of item appeared to be good discriminators. Table 22 gives the details and permits comparison between functional and content items.

FUNCTIONAL ITEMS :	DEIXIS	CONJ	PREP	PRO	AUX	ADV	SUB-TOTAL
ITEMS DELETED	17	10	11	6	4	3	= 51
NO.DISCRIMINATING	13	8	7	5	4	2	= 39

CONTENT ITEMS:	NOUN	VERB	ADJECTIVE	ADV	SUB-TOTAL	TOTAL
ITEMS DELETED	27	12	10	-	= 49	51F + 49C = 100
NO.DISCRIMINATING	14	7	4	-	= 25	39F + 25C = 64

TABLE 22 : ITEM ANALYSIS ACCORDING TO WORD CLASS

Abbreviations used: conj = conjunction, prep = preposition
pro = pronoun, aux = auxiliary
adv = adverb

These then were the results obtained from the British sample. Let us now turn briefly to report results obtained from the overseas samples.

8.1.1.2 **Overseas Samples** N = 85

The largest of the three samples was a group of 53 German students enrolled in various courses at Nuremberg University. They were receiving instruction through English. 46 of these had sat a 120 item Proficiency in English test (GERM.T.) administered at Christmas by the English Department of the University. The cloze test was given the following term, and results are presented below.

N = 53	MEAN	S.D.	SM	RELIABILITY	CORRELATION	
					CLOZE(Accep)	GERM.T.
CLOZE(Exact)	41·09	7·60	1·04	·68	·93*	·23
CLOZE (Acceptable)	54·40	10·86	1·49	·85	-	·28
GERM.T.(N=46)	44·53	6·42	·95	·45	-	-

TABLE 23 : RESULTS OF CLOZE TEST : NUREMBERG SAMPLE (N = 53)

Of the 100 items, only 35 discriminated, of which 20 were functional items and the remaining 15 represented content words. Much more interesting was the difference in correlation between the two systems of scoring i.e. verbatim and acceptable, when one compared upper, middle and lower groups:

Upper third : r = ·61* (N = 18)
Middle third : r = ·64* (N = 17)
Lower third : r = ·91* (N = 18)

Detailed analysis regarding FV and DI, along with individual results are contained in Appendix 4.C.2. and 4.C.4. The two other samples were both small, with only sixteen subjects each, one from Krefeld in W. Germany, and the other from Egypt. The results of the latter group were reported earlier in Experiment I (B), Section 7.2.1.2. but are repeated here for convenience. Because of their small size, no item analysis was done for either of these groups.

	MEAN	S.D.	SM	RELIABILITY	CORRELATION
CLOZE (Exact)	52.81	4.86	1.21	.2	.91*
CLOZE(Acceptable)	64.94	5.52	1.38	.38	

TABLE 24 : RESULTS OF CLOZE TEST : KREFELD SAMPLE (N = 16)

	MEAN	S.D.	SM	RELIABILITY	CORRELATION	
					CLOZE (Accep)	INST.T.
CLOZE (Exact)	46.88	8.5	2.13	.74	.98*	.3
CLOZE(Acceptable)	64.5	10.76	2.69	.84		.29

TABLE 25 : RESULTS OF CLOZE TEST : ALEXANDRIA SAMPLE (N = 16)

As shown there was no external criterion for the Krefeld sample, and with the Egyptian sample, correlation with their own test was low.

8.1.1.3 Summary

To briefly summarise the results of this part of the experiment; we found that this cloze test was a good discriminator between subjects with different levels of reading ability as evidenced by the very satisfactory reliability and validity coefficients obtained by using the

ELBA and EPTB as external criteria.

Results with overseas samples were less satisfactory primarily because of the small size of two of the samples used, and the absence of reliable external measures as comparitive criteria. Of interest is the similarity in mean between the two smaller overseas samples.

8.1.2 The Multiple-Choice Test : Evaluation of the Reading Product

This particular part of the experiment adopted procedures normally used in the development of objective type testing materials, that is, results of one try-out were used for development of the next version to be tested. Item retention was in question when the criteria adopted for the cloze test concerning FV and DI were applied here. However since the preparation of items was in accordance with a taxonomy of abilities considered desirable for NNS, retention of items was on occasion dictated by criteria other than that generally governed by satisfactory FV and DI. The latter were obtained using the E(1-3)method.

As with the cloze, the preliminary try-out was untimed and the time-limit of 45 minutes imposed on the first test version was considered adequate. The material used for this test was the same passage used for the cloze test. As the results of Experiment I (A) had established that cloze had little effect on the results of the multiple-choice test based on the same passage that followed after it, we present the results of the multiple-choice tests without

any attenuation of results due to ordering. The try-out version is included in Appendix 1.B.1.4.

A total of 250 students were involved in this part of the experiment and the results will be discussed as follows :

Version 1 - N = 60
Version 2 - N = 152
Version 3 - N = 38

8.1.2.1 Multiple-Choice Test : Version 1 (N = 60)

The results of this test are presented below with reliability and validity coefficients. For the latter, the ELBA total and the Reading Comprehension sub-test were once again used with a sample. Since Version 2 was based on this first experimental test, the item analysis will be discussed very briefly in this section rather than delayed till Chapter Nine.

The frequency distribution (Fig.15) shows the slight positive skewing. A test of significance revealed that the lower third group was significantly different from both the upper and middle third groups at ·05 level.

Upper and Lower third groups : F = 2·52
Upper and Middle third groups: F = 1·06
Middle and Lower third groups: F = 2·66
Value needed for significance at ·05 level was : F = 2·15

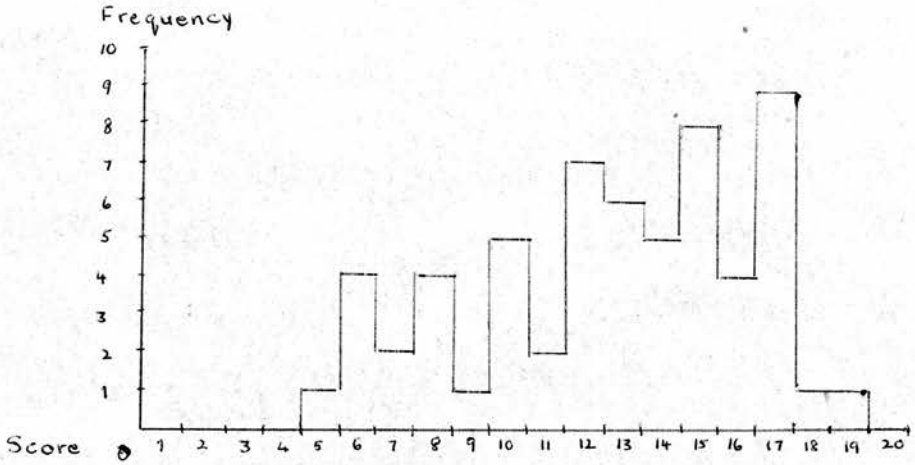


Fig. 15 FREQUENCY DISTRIBUTION : MULTIPLE-CHOICE VERSION 1
N = 60

Table 26 presents the results of the test along with reliability and validity coefficients.

	MEAN	S.D.	SM	RELIABILITY	CORRELATION ELBA(Total)	ELBA(RC)
MULTIPLE-CHOICE 1	12.70	3.62	.47	.75	.66*	.71*
ELBA(N=25)	77.64	14.39	2.88	.76	-	-

TABLE 26 : RESULTS OF MULTIPLE-CHOICE TEST : VERSION 1 (N=60)

The item analysis was made using the same criteria adopted for the cloze test in Section 8.1.1.1., which was:

F.V. to be .25 to .85

D.I, to be .25 and above

Table 27 gives the results. Thirteen of the twenty items satisfied the above FV and DI criteria but because some of the distractors had failed to function, marginal revisions were carried out to improve the quality of the items. Of the seven items that failed to discriminate adequately, all but one were revised and retained. This

decision was taken because it was felt by this writer that they represented items that reflected the taxonomy of abilities NNS were expected to need for reading comprehension at this level, and as such were necessary for the test. Items 1, 4, 9 and 17 had a very high facility value. Item 1 was retained since it was considered psychologically good to begin with an easy item. Item 17 asked subjects to identify the main theme and since this was one of the specific abilities listed, could not be discarded. Both these items were however revised as was item 4. Item 9 was the only item in Version 1 which was replaced. This asked for the meaning of a word in context and was easily replaced by a similar item requiring inference of meaning with distractors taken from the local context.

Item 8 on the other hand though satisfying the criteria, showed that more subjects preferred the wrong distractor. But in the opinion of this writer, the item seemed a fair test of inferential ability and was retained with slight revision. As it turned out this was <u>not</u> a wise decision and the item had eventually to be discarded though it was retained for Version 2.

Details of the item analysis are given below in Table 27.

ITEM NO.	DISTRACTOR (a)	(b)	(c)	(d)	OMITTED NO.	FACILITY VALUE	DISCRIMINATION INDEX	COMMENT	ADDITIONAL CLASSIFICATION
1	-	[53]	3	3	1	·88	·3	Revise	F. NL..-D
2	4	11	3	[39]	3	·65	·45	✓	I. NL. -D
3	[41]	2	6	9	2	··68	·6	✓	R. L. -D
4	2	-	4	[53]	1	·88	·25	Revise	I. L. -D
5	9	-	[38]	12	1	·63	·8	✓	R. L. -D
6	[51]	-	5	4	-	·85	·35	✓	F. NL. -D
7	3	5	[48]	2	2	·8	·4	✓	R. NL. -G
8	28	[16]	11	2	3	·27	·3	✓	I. L. -D
9	-	[56]	3	-	1	·93	·1	Replace*	I. L. -D
10	[50]	3	3	4	-	·83	·3	✓	R. L. -D
11	6	5	[41]	6	2	·68	·5	✓	I. NL. -G
12	[22]	12	11	13	2	·37	·7	✓	I. NL. -D
13	[15]	2	15	[26]	2	·43	·15	Revise	E. NL. -G
14	10	9	[32]	6	3	·53	·65	✓	E. NL. -G
15	22	[17]	1	14	6	·28	·4	✓	E. L. -D
16	1	[45]	10	2	2	·75	·4	✓	F. NL. -G
17	2	2	1	[54]	1	·9	·2	Revise	E. NL. -G
18	4	8	[45]	1	2	·75	·2	Revise	E. NL. -D
19	[13]	1	33	5	8	·22	·45	Revise	I. NL. -G
20	7	6	20	[22]	5	·37	·45	✓	E. NL. -G

TABLE 27 : ITEM ANALYSIS (N = 60) M.C. 1 [ ] = Correct answer

Key to additional classification used to reflect the taxonomy of abilities:

F = Factual R = Referential I = Inferential E = Evaluative

L = Localised distractors NL = non-localised distractors D = Detail

G = Generalized concepts

Since one of the hypotheses was aimed at investigating the effects of location of distractors on item difficulty, the items were classified into two groups to permit a comparison to be made. Table 28 presents the results of this analysis.

Items with localised distractors	Items with non-localised distractors
7 items : Score = 271	13 items : Score = 491
Mean = 4.52 or 64.52%	Mean = 8·18 or 62·95%

TABLE 28 : ANALYSIS ACCORDING TO LOCATION OF DISTRACTORS

Because of the unequal number of items, we adopted the following method. Using the results of all 60 subjects, we obtained the mean score based on the number of items involved. By reducing this to a percentage, we could compare performance. We did a similar analysis for the items based on our taxonomy of abilities.

Factual items: 1, 6, 16 = 3 Total score = 149 Mean = 2·48 or 82·78%
Referential items: 3, 5, 7, 10 = 4 Total score = 177 Mean = 2·95 or 73·75%
Inferential items: 2, 4, 8, 9,11,12,19 = 7 Total score = 240 Mean = 4 or 57·14%
Evaluative items: 13,14,15,17,18,20 = 6 Total score = 196 Mean = 3·27 or 54·44%

TABLE 29 : ANALYSIS ACCORDING TO TAXONOMY

The results showed a tendency for difficulty to increase as subjects attempted questions aimed at testing inferential and evaluative items. We recognised that very rarely was an item testing only one ability, but the classification was made according to the experimenter's opinion regarding the primary function of the questions set.

Version 2 was developed on the basis of the results just reported. A distribution of the questions in terms of the abilities being tested is given in Appendix 1.B.1.5. for all three versions of the multiple-choice test. Copies of the different versions developed for use are contained in Appendices 1.B.1.6., 1.B.1.7, and 1.B.1.8.

8.1.2.2 Multiple-Choice Test Version 2 (N = 152)

This second revised version was tried out with the largest sample and is described in greater detail since it was this version that was largely used for the third part of this experiment i.e. the comparison of the process and product via the cloze and multiple-choice tests.

Of the total of 152 students who took the test in Version 2, 97 came from the British sample and 55 from the two German samples in Nuremberg and Krefeld. Assuming a null hypothesis we carried out an analysis of variance which confirmed that all three samples were indeed drawn from the same overall population enabling us to combine the data for item analysis. Results of Version 2 then are based on the total sample who sat this test.

	Number	Sum of Scores	Sum of Scores Squared
British sample	97	1008	11586
Nuremberg sample	39	455	5575
Krefeld sample	16	185	2217
Total	152	1648	19,378

	SUM OF SQUARES	df	MEAN SQUARE	F-VALUE
A BETWEEN GROUPS	54·49295632	(3-1) =2	A 27·24647816	A ÷ B =
B WITHIN GROUPS	1455·71757	(152-3) = 149	9·769916577	2.788813798
	1510·210526	151		

TABLE 30 : ANALYSIS OF VARIANCE

A value of F = 3·69 was needed for significance at the ·05 level, hence we accepted the null hypothesis that the three samples were all representative of the same parent population.

The frequency distribution of the scores bears out this assumption. In Fig. 16, the total sample distribution has been superimposed on the frequency distribution of the British sample.

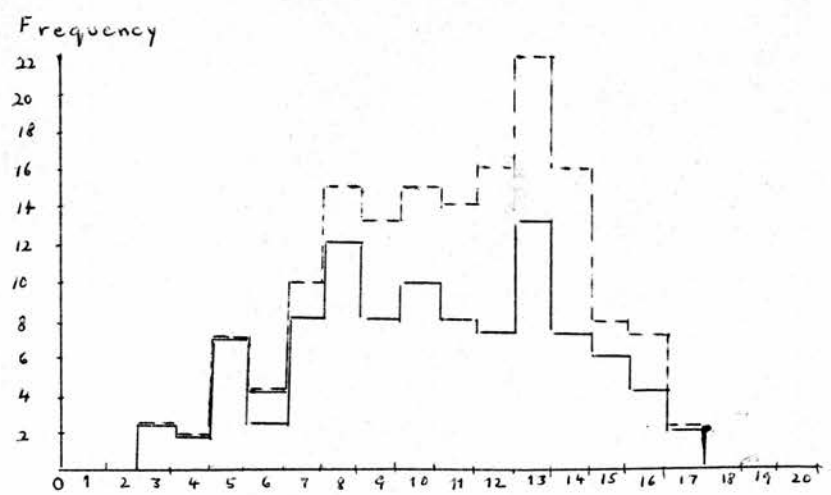


Fig. 16. FREQUENCY DISTRIBUTION : MULTIPLE-CHOICE VERSION 2

—— British sample (N = 97)

--- Total sample (N = 152)

The results of all 152 subjects are presented in Table 31, along with reliability and validity coefficients. For the latter, the EPTB was used as a criterion measure. In addition two other sub-groups of the sample had sat internal tests of language proficiency. These are identified as the GERM.T. and the LANG.T. The GERM.T. was administered by the English Department of Nuremberg University to their students about two months prior to this experiment. It was the same 120-item test reported in 8.1.1.2. The LANG.T. referred to in 8.1.1.1. was, as reported, administered by a language school in Britain.

	MEAN	S.D.	SM	RELIABILITY	CORRELATION		
					EPTB	GERM.T.	LANG.T
MULTIPLE-CHOICE	10·84	3·16	.26	·65	·5*	·13	·27
EPTB (N = 42)	34·73	4·23	·65	-	-	-	-
GERMAN TEST (N = 37)	44·81	5·95	·1	·36	-	-	-
LANGUAGE TEST (N = 34)	67·29	13·47	2·31	Not available	-	-	-

TABLE 31 : RESULTS OF MULTIPLE-CHOICE : VERSION 2 (N = 152)

Item analysis details will be reviewed in Chapter Nine, but a few comments are presented here since they are partially responsible for some of the changes in Version 3 of the test. For this second version, sixteen of the twenty items showed a satisfactory facility value and discrimination index. Most of the revised items showed an improvement in quality. With the larger sample, none of the items was too easy, but three had a poor discrimination value, and one item, No.13, had an FV of less than .25. Table 32 presents the detailed analysis.

ITEM NO.	DISTRACTOR (a)	(b)	(c)	(d)	OMITTED NO.	FACILITY VALUE	DISCRIMINATION INDEX	COMMENT	ADDITIONAL CLASSIFICATION
1	15	[118]	14	1	4	·78	·2	Retain*	F. NL. -D
2	12	24	10	[100]	6	·66	·35	✓	I. NL. -D
3	[88]	6	10	47	1	·58	·37	✓	R. L. -D
4	8	5	[77]	55	7	·51	·43	✓	I. L. -D
5	25	3	[86]	37	1	·57	·27	✓	R. L. -D
6	[120]	6	14	11	1	·79	·27	✓	F. NL. -D
7	2	18	[113]	6	13	·74	·41	✓	R. NL. -G
8	39	[40]	56	4	13	·26	·24	Replace*	I. L. D
9	14	38	14	[80]	6	·53	·37	✓	I. L. D
10	[125]	2	9	14	2	·82	·27	✓	R. L. D
11	30	22	[52]	36	12	·34	·53	✓	I. NL. G
12	25	32	16	[68]	11	·45	·37	✓	I. NL. D
13	[32]	17	44	43	16	·21	·31	Revise	E. NL. G
14	20	18	[70]	30	14	·46	·45	✓	E. NL. G
15	55	[44]	8	24	21	·29	· 2	Revise	E. L. D
16	12	[84]	24	15	17	·55	·43	✓	F. NL. G
17	3	2	14	[124]	9	·82	·25	✓	E. NL. G
18	27	10	[84]	17	14	·55	·37	✓	E. NL. D
19	[83]	7	26	9	27	·55	·43	✓	I. NL. G
20	25	13	34	[60]	20	·39	·43	✓	E. NL. G

TABLE 32 : ITEM ANALYSIS (N = 152) M.C. [ ] Correct answer

Item 1 was to be retained although DI was low. As reported earlier Item 8 was replaced because the wrong distractor continued to attract most of the attention despite revision. Items 13 and 15 were revised because both showed that other distractors were drawing the lion's share of attention.

In addition to this item analysis, two other analyses were done as for Version 1. The first was concerned with the effects of location of distractors on item difficulty, and the second was performed on the basis of the taxonomy. The additional comment in Table 32 shows how the items were distributed, the abbreviations used being the same as explained in Table 27. Table 33 presents the results of both types of analyses for the group as a whole.

Items with localised distractors 7 items: Score = 540 Mean = 3·55 or 50·75%	Items with non-localised distractors 13 items: Score = 1108 Mean = 7·29 or 56·07%
FACTUAL ITEMS: 1,6,16 = 3 Total Score = 322, Mean = 2·12 or 70·61%	
REFERENTIAL ITEMS: 3,5,7,10 = 4 Total Score = 412, Mean = 2·17 or 67·76%	
INFERENTIAL ITEMS: 2,4,8,9,11,12,19 = 7 Total Score = 500, Mean = 3·29 or 46·99%	
EVALUATIVE ITEMS: 13,14,15,17,18,20 = 6 Total Score = 414, Mean 2·72 or 45·39%	

TABLE 33 : ANALYSIS OF RESULTS ACCORDING TO TAXONOMY: VERSION 2

Since this was the version used with the largest sample, a further detailed analysis for each of the upper, middle and lower thirds was carried out to enable a comparison of

performance to be made. This is given in Table 34. As with the earlier experiments, individual results are given in Appendix 4.C.3. and 4.C.4.

	UPPER THIRD N=51	MIDDLE THIRD N=50	LOWER THIRD N=51
LOCALISED DISTRACTORS No = 7	Score = 244 $\bar{x}$=4.78 or 68.35%	Score = 162 $\bar{x}$=3.24 or 46.29%	Score = 134 $\bar{x}$=2.63 or 37.54%
NON-LOCALISED DISTRACTORS No = 13	Score = 480 $\bar{x}$=9.41 or 72.4%	Score = 395 $\bar{x}$=7.9 or 60.77%	Score = 233 $\bar{x}$=4.57 or 35.14%
FACTUAL ITEMS:3	Score = 131 $\bar{x}$=2.57 or 85.62%	Score = 106 $\bar{x}$=2.12 or 70.67%	Score = 85 $\bar{x}$=1.67 or 55.56%
REFERENTIAL ITEMS:4	Score = 168 $\bar{x}$=3.29 or 82.35%	Score = 144 $\bar{x}$=2.88 or 72%	Score = 100 $\bar{x}$=1.96 or 49.02%
INFERENTIAL ITEMS:7	Score = 244 $\bar{x}$=4.78 or 68.35%	Score = 151 $\bar{x}$=3.02 or 43.14%	Score = 105 $\bar{x}$=2.06 or 29.41%
EVALUATIVE ITEMS:6	Score = 181 $\bar{x}$=3.55 or 59.15%	Score = 156 $\bar{x}$=3.12 or 52%	Score = 77 $\bar{x}$=1.51 or 25.16%

TABLE 34 : RESULTS OF MULTIPLE-CHOICE : VERSION 2 (N = 152)

($\bar{x}$ = Mean)

As in Version 1, there was no difference between items with localised and non-localised distractors with upper and lower thirds, although with the items divided according to the taxonomy, there was the same trend noticed in the earlier results. The factual and referential items were markedly easier than inferential and evaluative items for all, but more so with middle and lower thirds. Let us now briefly report the results of the third version

which introduced seven new items, five concerned with textual structure and organization. Of the remaining two, one replaced Item 8, and another tried to judge the reader's sensitivity to the author's style. To accommodate the seventh item, Item 6, a good item was replaced.

8.1.2.3 Multiple-Choice Test : Version 3

This version had one main aim, the addition of items that attempted to test whether textual organization was comprehended and its value appreciated in use. Of the three items that had failed the criteria in Version 2, two were revised. But the interest here lay in finding out whether the textual structure was an area worth investigating. This version was tried out with a total of 38 students drawn from two sources, one British sample and the other a group of Egyptian students in Alexandria. The latter group was referred to earlier in Experiment I (B) where the multiple-choice and cloze tests were sat in counter-balanced order. As already stated the cloze results did not reveal significant differences due to ordering, but the multiple-choice did. Because of this, the results were not analysed, so that the results of 22 students who constituted the British sample, are the only evidence reported here and in Appendix 4.C.3. However because the sample was too small for us to utilise the E_{1-3} formula for item discrimination, we adopted what Downie and Heath (1965:232) recommended in such cases. They argued that an effective though crude measure would be a difference of at least two between

the frequency of high and low groups scoring the item successfully, if the item was to be regarded as discriminating.

There was no external validity criterion for this group apart from the cloze test prepared by this writer. Results of this correlation will however be reported later when a test of process is evaluated against a test of the end-product of reading.

The frequency distribution, mean, standard deviation and item analysis are reported below. There were 25 items although the time limit of 45 minutes was not extended. As results indicate, the new items were not particularly successful, and some of the previously tested items, e.g. Items 7 and 10, failed to qualify with this sample. Details will be discussed in Chapter Nine.

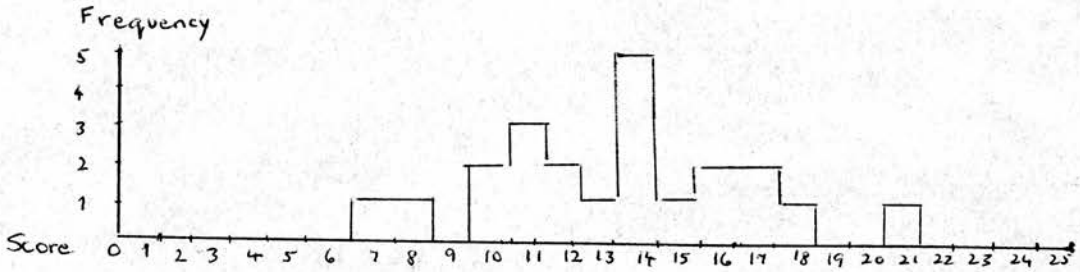


Fig. 17. FREQUENCY DISTRIBUTION : MULTIPLE-CHOICE VERSION 3 (N = 22)

	MEAN	S.D.	SM	RELIABILITY
MULTIPLE-CHOICE - 3	13·41	3·36	·72	·60

TABLE 35 : RESULTS OF MULTIPLE-CHOICE : VERSION 3 (N = 22)

ITEM NO	DISTRACTOR (a)	(b)	(c)	(d)	OMITTED NO	FACILITY VALUE	DISCRIMINATION INDEX	COMMENT	ADDITIONAL CLASSIFICATION
1	-	[21]	1	-	-	·95	·14	Retain	F. NL. -D
2	1	5	-	[15]	1	·68	·29	✓	I. NL. -D
3	[13]	1	-	8	-	·59	·29	✓	R. L. -D
4	4	1	[17]	3	-	·77	·29	✓	I. L. -D
5	4	-	[12]	6	-	·55	·29	✓	R. L. -D
6	[3]	9	5	4	1	·14	·14	New	E. L. -G
7	-	2	[19]	-	1	.86	.43		R. NL. -G
8	4	4	11	[0]	3	0	0	New	I. L. -D
9	3	2	1	[15]	1	·68	·71	✓	I. L. -D
10	[18]	-	1	2	1	·82	·14		R. L. -D
11	6	-	[13]	2	1	·59	·57	✓	I. NL. -G
12	[7]	1	8	4	2	·32	·43	✓	I. NL. -D
13	-	3	[13]	4	2	·59	·43	✓	E. NL. -G
14	3	4	-	[11]	4	·50	·29	✓	E. NL. -G
15	1	[6]	4	7	4	·27	·57	✓	E. L. -D
16	1	[15]	4	1	1	·68	·29	✓	F. NL. -G
17	-	-	1	[20]	1	·91	0	Retain	E. NL. -G
18	5	-	[12]	4	1	·55	·43	✓	E. NL. -D
19	[13]	1	-	-	8	·59	·57	✓	I. NL. -G
20	7	3	3	[7]	2	·32	·57	✓	E. NL. -G
21	[8]	4	5	2	3	·36	·29	New ✓	E. NL. -G
22	4	[7]	2	6	3	·32	·14	New	E. NL. -G
23	12	[4]	5	-	1	·18	·29	New	E. NL. -G
24	2	2	[14]	1	3	·64	·29	New ✓	E. L. -D
25	5	[12]	-	-	5	.56	.43	New ✓	E. NL. -G

TABLE 36 : ITEM ANALYSIS M.C. 3 [ ] Correct answer

Note: 'New' refers to the items introduced in this version. The abbreviations used in the 'additional classification' column are as for the two earlier versions.

An analysis of the results based on the taxonomy, and classified according to the type of distractors used for the item was carried out similar to that done for Versions 1 and 2. The results are presented below in brief.

Items with localised distractors: 9 items
Score = 98, Mean = 4.45 or 49.44%

Items with non-localised distractors: 16 items
Score = 197, Mean = 8.95 or 55.96%

Factual items: 2 items
Score = 36, Mean = 1.64 or 81.83%

Referential items: 4 items
Score = 62, Mean = 2.82 or 70.20%

Inferential items: 7 items
Score = 80, Mean = 3.64 or 51.95%

Evaluative items: 12 items
Score = 117, Mean = 5.32 or 44.32%

As we can see, localisation of distractors appeared to have little or no effect on item difficulty. Success with items classified according to the taxonomy followed the same trend exhibited by subjects in the two earlier versions. Factual and referential items were much easier than inferential and evaluative items. Five of the latter were new and were primarily concerned with testing whether subjects were sensitive to text structure and organization, but as the item analysis demonstrated were not particularly good as items.

8.1.2.4 Summary

The multiple-choice test was for the first two versions based on abilities we needed to evaluate and which

had been listed in the taxonomy specified. The third version was primarily aimed at 'other' skills associated with textual structure and organisation. We come now to the third part of the experiment.

8.1.3 Evaluation of Process Versus Product

The reliability and validity of the two tests used to measure the process and product had to be established if this part of the experiment was to have any validity. Having done so, we correlated the two measures which were viewed as providing concurrent validity. The samples are reported according to the three versions of the test of the reading product.

Cloze versus Multiple-choice - Version 1 : N = 25

Cloze versus Multiple-choice - Version 2 : N = 152

Cloze versus Multiple-choice - Version 3 : N = 22

However since the results of acceptable cloze for the Version 2 sample of 152 composed of 97 students from the British sample, and 55 German students (39 from Nuremberg and 16 from Krefeld), showed that there were significant differences, results for these three groups will be presented independently. In each instance cloze preceeded the multiple-choice test.

Each correlation is accompanied by details of reliability and external validity criteria where these were available.

8.1.3.1 Cloze Versus Multiple-Choice: Version 1 (N = 25)

N = 25	MEAN	S.D.	SM	REL.	CORRELATION		
					CLOZE (Accep)	MC-1	ELBA(N=16
CLOZE(Exact)	50.2	10.72	2.14	·84	·91*	·75*	·58*
CLOZE (Acceptable)	69.48	14.03	2.81	·91	-	·82*	·69*
MULTIPLE-CHOICE 1	14.44	3.27	·65	·68		-	·73*
ELBA(N=16)	84.13	9.6	2.4	·45			-

TABLE 37 : CONCURRENT VALIDITY OF CLOZE & MULTIPLE-CHOICE 1

8.1.3.2 Cloze Versus Multiple-Choice: Version 2 (N = 152)

N = 97	MEAN	S.D.	SM	REL.	CORRELATIONS			
					CLOZE (Accep)	MC-2	EPTB	LANG.T.
CLOZE (Exact)	27.35	10.85	1.1	.85	·98*	·45*	·51*	·47*
CLOZE (Accep)	35.16	13.99	1.42	·91	-	·49*	·59*	·52*
MULTIPLE-CHOICE 2	10.39	3.4	·35	·71		-	·5*	·27
EPTB (N=42)	34.73	4.23	·65	-			-	-
LANGUAGE TEST (N=34)	67.29	13.47	2.31	-			-	-

TABLE 38 : CONCURRENT VALIDITY OF CLOZE & MULTIPLE-CHOICE 2

BRITISH SAMPLE (N = 97)

N = 39	MEAN	S.D.	SM	REL	CORRELATION		
					CLOZE (Accep)	MC-2	GERMAN.T.
CLOZE(Exact)	41.85	7.31	1.17	·65	·92*	·45*	·17
CLOZE (Acceptable)	55.26	10.03	1.61	·80	-	·53*	·21
MULTIPLE-CHOICE-2	11.67	2.65	·42	·48		-	·13
GERMAN TEST (N = 37)	44.81	5.95	·98	·36			-

TABLE 39 : CONCURRENT VALIDITY OF CLOZE & MULTIPLE-CHOICE 2

NUREMBERG SAMPLE (N = 39)

There was an interval of six weeks between the cloze and the multiple-choice test for the Nuremberg sample. The 'German' test was, as reported earlier, an English test administered by the English Department of the University, and despite including 120 items had very low reliability.

For the third group, no external criterion was available.

This was the sample from Krefeld.

N = 16	MEAN	S.D.	SM	RELIABILITY	CORRELATION	
					CLOZE (Accep)	MC 2
CLOZE(Exact)	52.81	4.86	1.21	·2	·91*	·43
CLOZE (Acceptable)	64.94	5.52	1.38	· 38	-	·46
MULTIPLE-CHOICE-2	11.56	2.28	·57	·28		-

TABLE 40 : CONCURRENT VALIDITY OF CLOZE & MULTIPLE-CHOICE 2

KREFELD SAMPLE (N = 16)

The correlations were not significant and could have occurred by chance at marginally above the five percent level. With these smaller samples, reliability was much less satisfactory as compared to the British Sample.

8.1.3.3 Cloze Versus Multiple-Choice - Version 3 (N = 22)

N = 22	MEAN	S.D.	SM	RELIABILITY	CORRELATION	
					CLOZE (Accep)	MC-3
CLOZE(Exact)	39.14	10.3	2.2	.83	.96*	.75*
CLOZE (Acceptable)	52.59	14.58	3.11	.92	-	.75*
MULTIPLE-CHOICE-3	13.41	3.36	.72	.60		-

TABLE 41 : CONCURRENT VALIDITY OF CLOZE & MULTIPLE-CHOICE 3

8.1.4 Summary of Experiment III Results

The three parts of this experiment involved the analysis of a large amount of data both because of the techniques adopted and because of the fairly large samples used. As our analysis showed, although there was a positive correlation between the process and product stages, often significant at more than 5% level, the shared variance was far from unity. The implications will be discussed in Chapter Nine. We come now to Experiment IV.

8.2 EXPERIMENT IV : INVESTIGATION OF LENGTH AS VARIABLE

This experiment was primarily designed to investigate how students of varying levels of proficiency in the second language reacted to different lengths of text again at first

exposure. Only two lengths were investigated permitting a crude comparison but sufficient for our purposes. As our description in Chapter Six had indicated, four pairs of texts were used in this experiment, each pair drawn from a common source, written by the same author on the same topic. Each pair comprised a long and a short text presented in counter-balanced order to cancel out possible effects of ordering on performance. All texts were presented as cloze tests with the 'any' word deletion method where every tenth word was omitted and replaced by a standard underlined length of 14 spaces. The first and last sentence of each text was left unmutilated. All responses were scored by exact and acceptable word methods. Responses that did not satisfy criteria for acceptability were marked incorrect. Unmarked items were scored as omitted items. Copies of tests and instructions are given in Appendix 1, while the results for each pair of texts is contained in Appendix 4.D.

Each pair of texts was presented to a different sample. There were therefore four samples drawn mainly from the same overall population used in Experiment III and all subjects except two, had sat the Cloze test Passage 1 reported in the earlier experiment. These subjects were students studying in Britain. Instructions were presented orally as well as printed in the test booklets.

Results of each sample are presented independently. In addition to means, standard deviations and reliability coefficients, correlation coefficient between the short and

long texts are presented. Also provided is correlation coefficients of each text with the Cloze Test - Passage 1 common to all subjects. Since two scoring procedures were used, comparisons are reported for each of the methods.

Item analysis was carried out and the criteria adopted for assessing items as discriminating were as for the experiments reported earlier : an F.V. of ·25 to ·85 and a D.I. of ·25 and above. Items were also analysed according to the word class they represented in the context of use. We are reminded that all A passages were long, and B passages short.

8.2.1 Passage 2 What is Law? (N = 49)

N = 49	MEAN	S.D.	SM	RELIABILITY	NO. OF ITEMS
P 2A (Exact)	32.24	9.28	1.33	·79	100
P 2A(Acceptable)	40.49	12.53	1.79	·89	100
P 2B (Exact)	5.29	1.85	·26	·41	25
P 2B(Acceptable)	6.73	2.76	·39	·39	25
P 1 (Exact)	28.31	9.53	1.36	·80	100
P 1 (Acceptable)	36.71	12.65	1.81	·89	100

TABLE 42 : RESULTS OF CLOZE TEST, PASSAGE 2

N = 49	P2A (Accep)	P2B (Exact)	P2B (Accep)	P1 (Exact)	P1 (Accep)
P 2A (Exact)	·98*	·52*	-	·81*	-
P 2A(Acceptable)	-	-	·69*	-	·85*
P 2B (Exact)		-	·89*	·47*	-
P 2B(Acceptable)			-	-	·64*
P 1 (Exact)				-	·97*
P 1 (Acceptable)					-

TABLE 43 : CORRELATION COEFFICIENTS : P2A, 2B & P1

The range of scores by acceptable word method varied from 20 to 68 in 2A, and from 1 to 12 in 2B. An item analysis was carried out to determine facility value and discrimination index of each item. Details are contained in Appendix 4.D.3.1. Only 34 items succeeded in discriminating between upper and lower thirds in the long passage, while 7 items out of 25 were efficient in the short text. Both passages were therefore judged to be inefficient. The discriminating items were then classified according to word class and the results are presented below in Tables 44 and 45.

FUNCTIONAL ITEMS:	DEIXIS	CONJ	PREP	PRO	AUX	ADV	SUB-TOTAL
ITEMS DELETED	15	13	6	4	8	3	= 49
NO.DISCRIMINATING	7	5	1	2	4	1	= 20

CONTENT ITEMS:	NOUN	VERB	ADJECTIVE	SUB-TOTAL	TOTAL
ITEMS DELETED	20	21	10	= 51	49F + 51C = 100
NO.DISCRIMINATING	8	5	1	= 14	20F + 14C = 34

TABLE 44 : ITEM ANALYSIS ACCORDING TO WORD CLASS P 2A

FUNCTIONAL ITEMS:	DEIXIS	CONJ	PREP	PRO	ADV	SUB-TOTAL
ITEMS DELETED	1	2	5	1	4	= 13
NO.DISCRIMINATING	-	-	4	-	-	= 4

CONTENT ITEMS:	NOUN	VERB	ADJECTIVE	SUB-TOTAL	TOTAL
ITEMS DELETED	6	5	1	= 12	13F + 12C = 25
NO.DISCRIMINATING	-	3	-	= 3	4F + 3C = 7

TABLE 45 : ITEM ANALYSIS ACCORDING TO WORD CLASS - P2B

A similar analysis is presented for Passages 3, 4 and 5 now.

8.2.2 PASSAGE 3 URBANISATION (N = 36)

N = 36	MEAN	S.D.	SM	RELIABILITY	ITEMS
P 3A (Exact)	29.36	13.13	2.19	.90	100
P 3A(Acceptable)	37.53	17.01	2.84	.94	100
P 3B (Exact)	8.22	3.98	.66	.73	25
P 3B(Acceptable)	10	4.68	.78	.82	25
P 1 (Exact)	26.58	13.16	2.19	.90	100
P 1 (Acceptable)	34.28	16.48	2.75	.94	100

TABLE 46 : RESULTS OF CLOZE TEST, PASSAGE 3

Acceptable Score Range: 9 to 72 in 3A, and 1 to 18 in 3B

	P3A(Accep)	P3B(Exact)	P3B(Accep)	P1(Exact)	P1(Acce
P 3A (Exact)	.99*	.80*	-	.87*	-
P 3A(Acceptable)	-	-	.83*	-	.89*
P 3B (Exact)		-	.97*	.78*	-
P 3B(Acceptable)			-	-	.82*
P 1 (Exact)				-	.99*
P 1 (Acceptable)					-

TABLE 47 : CORRELATION COEFFICIENTS P3A, 3B & P1

FUNCTIONAL ITEMS:	DEIXIS	CONJ	PREP	PRO	AUX	ADV	SUB-TOTAL
ITEMS DELETED	9	3	15	4	2	4	= 37
NO.DISCRIMINATING	5	1	14	4	1	4	= 29

CONTENT ITEMS:	NOUN	VERB	ADJECTIVE	ADV	SUB-TOTAL	TOTAL
ITEMS DELETED	35	11	16	1	= 63	37F + 63C = 100
NO.DISCRIMINATING	15	6	6	-	= 27	29F + 27C = 56

TABLE 48 : ITEM ANALYSIS ACCORDING TO WORD CLASS P3A

FUNCTIONAL ITEMS:	DEIXIS	CONJ	PREP	AUX	SUB-TOTAL
ITEMS DELETED	5	2	3	1	= 11
NO.DISCRIMINATING	5	1	3	-	= 9

CONTENT ITEMS:	NOUN	VERB	ADJ	ADV	SUB-TOTAL	TOTAL
ITEMS DELETED	7	2	4	1	= 14	11F + 14C = 25
NO.DISCRIMINATING	2	1	3	1	= 7	9F + 7C = 16

TABLE 49 : ITEM ANALYSIS ACCORDING TO WORD CLASS P3B

A sample of 36 subjects sat the tests, and as Tables 48 and 49 show, there were 56 items in the long and 16 items in the shorter test that discriminated between upper and lower thirds. All correlations were performed with data from the same system of scoring. Details regarding item analysis are contained in Appendix 4.D.3.2. while individual results are reported on page 196 of Appendix 4.D.4.2.

8.2.3 Passage 4 The Leatherback Turtle N = 16

N = 16	MEAN	S.D.	SM	RELIABILITY	ITEMS
P 4A (Exact)	51.5	12.46	3.12	.87	115
P 4A(Acceptable)	75.44	14.45	3.61	.90	115
P 4B (Exact)	8.38	3.32	.83	.59	25
P 4B(Acceptable)	15.63	4.35	1.09	.78	25
P 1 (Exact)	52.14	5.56	1.49	.39	100
P 1 (Acceptable)	66.29	6.58	1.76	.57	100

TABLE 50 : RESULTS OF CLOZE TEST, PASSAGE 4

Acceptable Score Range: 5 to 104 in 4A, and 7 to 22 in 4B

	P4A (Accep)	P4B (Exact)	P4B (Accep)	P1 (Exact)	P1 (Accep)
P 4A (Exact)	.96*	.72*	-	.21	-
P 4A(Acceptable)	-	-	.79*	-	.72*
P 4B (Exact)		-	.91*	.25	-
P 4B(Acceptable)			-	-	.47
P 1 (Exact)				-	.82*
P 1 (Acceptable)					-

TABLE 51 : CORRELATION COEFFICIENTS P. 4A, 4B & P1

FUNCTIONAL ITEMS:	DEIXIS	CONJ	PREP	PRO	AUX	SUB-TOTAL
ITEMS DELETED	10	9	24	7	5	= 55
NO.DISCRIMINATING	1	4	7	3	3	= 18

CONTENT ITEMS:	NOUN	VERB	ADJ	ADV	SUB-TOTAL	TOTAL
ITEMS DELETED	27	18	11	4	= 60	55F + 60C = 115
NO.DISCRIMINATING	11	6	6	3	= 26	18F + 26C = 44

TABLE 52 : ITEM ANALYSIS ACCORDING TO WORD CLASS P.4A

FUNCTIONAL ITEMS:	DEIXIS	CONJ	PREP	PRO	AUX	ADV	SUB-TOTAL
ITEMS DELETED	1	1	2	2	1	1	= 8
NO.DISCRIMINATING	-	-	-	1	-	1	= 2

CONTENT ITEMS:	NOUN	VERB	ADJ	ADV	SUB-TOTAL	TOTAL
ITEMS DELETED	4	7	5	1	= 17	8F + 17C = 25
NO.DISCRIMINATING	2	6	1	-	= 9	2F + 9C = 11

TABLE 53 : ITEM ANALYSIS ACCORDING TO WORD CLASS P.4B

For both this and Passage 5, the samples were small, and the criteria for discrimination was the same as for Multiple-Choice Version 3 reported in Section 8.1.2.3. The item was classified as 'discriminating' if the facility value lay between ·25 and ·85, and there was a difference of at least 2 scores between upper and lower thirds.

P4A was the longest text with a total of 1203 words and a deletion count of 115 items. As results indicate this test was easy for the sample as a whole so that the number of effective items was disappointingly small, only 44. There were 11 effective items in the shorter test. Appendix 4.D.3.3. (page 185) gives details of item analysis, while Appendix 4.D.4.3 (page 198) gives the individual scores.

8.2.4 Passage 5 The Scientific Study of Social Development(N=16)

N = 16	MEAN	S.D.	SM	RELIABILITY	ITEMS
P 5A (Exact)	30.56	8.64	2.16	.75	100
P 5A(Acceptable)	46.13	13.85	3.46	.91	100
P 5B (Exact)	6.31	2.89	.72	.45	25
P 5B(Acceptable)	9.25	4.30	1.07	.77	25
P 1 (Exact)	39.38	11.42	2.85	.52	100
P 1(Acceptable)	53.44	16.33	4.08	.94	100

TABLE 54 : RESULTS OF CLOZE TEST, PASSAGE 5

Acceptable Score Range: 20 to 72 in 5A and 2 to 15 in 5B

	5A (Accep)	5B (Exact)	5B (Accep)	P1 (Exact)	P1 (Accep)
5A (Exact)	.94*	.70*	-	.83*	-
5A(Acceptable)	-	-	.78*	-	.77*
5B (Exact)		-	.96*	.61*	-
5B(Acceptable)			-	-	.72*
P1 (Exact)				-	.97*
P1(Acceptable)					-

TABLE 55 : CORRELATION COEFFICIENTS P.5A, 5B & P1

FUNCTIONAL ITEMS:	DEIXIS	CONJ	PREP	PRO	AUX	ADV	SUB-TOTAL
ITEMS DELETED	12	6	14	10	2	2	= 46
NO.DISCRIMINATING	4	4	3	4	2	1	= 18

CONTENT ITEMS:	NOUN	VERB	ADJ	ADV	SUB-TOTAL	TOTAL
ITEMS DELETED	26	15	11	2	= 54	46F + 54C = 100
NO.DISCRIMINATING	16	9	4	-	= 29	18F + 29C = 47

TABLE 56 : ITEM ANALYSIS ACCORDING TO WORD CLASS P.5A

FUNCTIONAL ITEMS:	DEIXIS	CONJ	PREP	PRO	ADV	SUB-TOTAL
ITEMS DELETED	3	2	2	2	1	= 10
NO.DISCRIMINATING	3	1	1	1	1	= 7
CONTENT ITEMS:	NOUN	VERB	ADJ	SUB-TOTAL		TOTAL
ITEMS DELETED	8	4	3	= 15		10F + 15C = 25
NO.DISCRIMINATING	2	-	2	= 4		7F + 4C = 11

TABLE 57 : ITEM ANALYSIS ACCORDING TO WORD CLASS P.5B

These tests with P5 had lower mean scores than P4. P5A proved to be a better discriminator between ability levels, with 47 effective items out of a possible 100 in Text A, but P5B had the same number of effective items as P4B, only 11. See Appendix 4.D.3.4. and 4.D.4.4. for details of performance.

8.2.5 General Analysis of Results

All four pairs of texts were analysed further by using the same technique adopted in 8.1.1.1. Each text was divided into blocks of ten items each for the long passages, and blocks of five items each for the shorter passages. Once again mean scores were calculated for upper, middle and lower third groups for each of the three categories : correct by acceptable standards, incorrect and omitted. These mean scores are contained in Appendix 4.D.2. but were translated into graphs which are presented here to permit comparison between upper and lower third groups for each of the eight tests (See pages 282 to 286)

Although the results will be discussed in the next chapter, a summary of the results of this experiment presented in Table 58 reveals that in general, longer texts were better predictors of performance for other long texts. It was much more difficult to predict performance across variations in length particularly when the texts were on different topics. For the table we are reminded that all A texts are long, and B texts are short. P1 was a long passage.

TEXT	P2B	P3B	P4B	P5B	P1
P2A	·69*				·85*
P3A		·83*			·89*
P4A			·79*		·72*
P5A				·78*	·77*
P1	·64*	·82*	·47	·72*	-

TABLE 58 : PREDICTIVE VALIDITY COEFFICIENTS FOR TEXTS OF DIFFERENT LENGTH (Scored by acceptable word method)

Figures 18 to 41 that follow present the analysis of the total response scores for all eight tests used in this experiment.

CORRECT RESPONSE SCORE : Acceptable Word Method

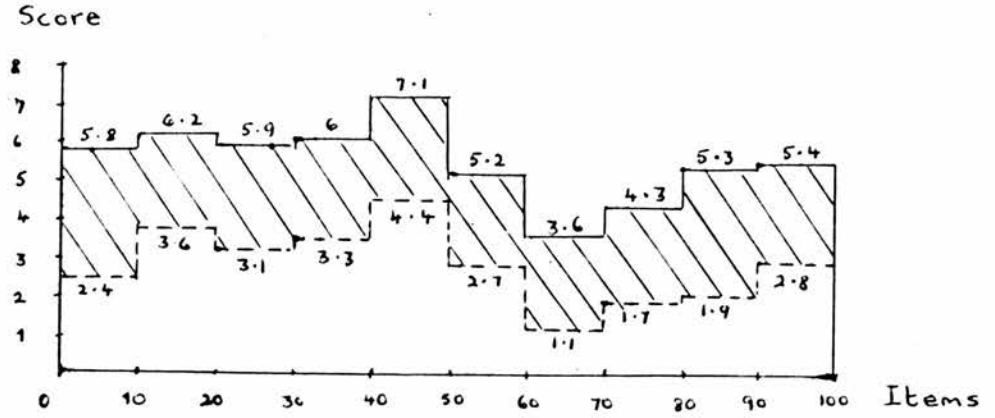


Fig. 18 PASSAGE 2A N = 49 —— Upper Third Mean = 54·81
--- Lower Third Mean = 26·94

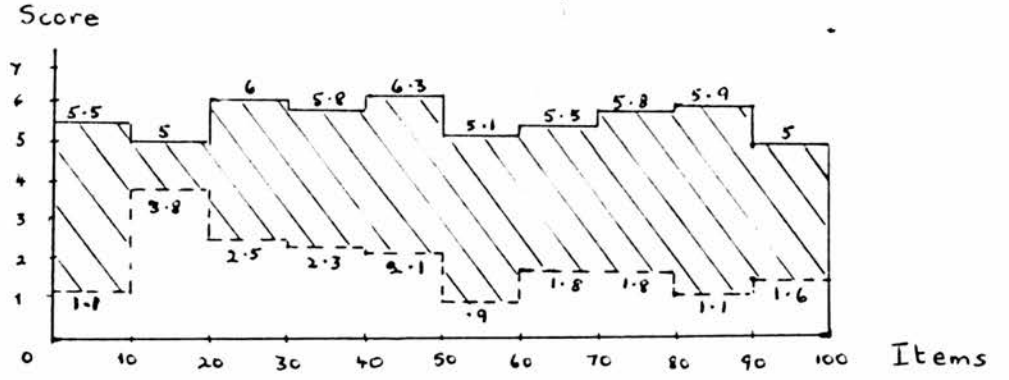


Fig. 19 PASSAGE 3A N = 36 —— Upper Third Mean = 55·92
--- Lower Third Mean = 18·75

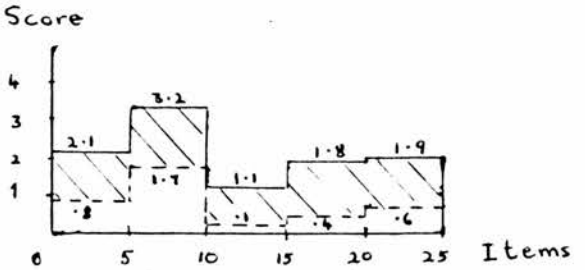


Fig. 20 PASSAGE 2B N = 49

—— Upper Third Mean = 9·94
--- Lower Third Mean = 3·63

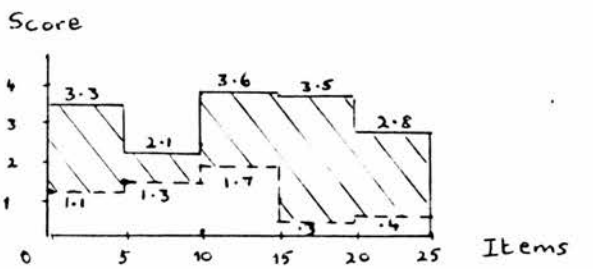


Fig. 21 PASSAGE 3B N = 36

—— Upper Third Mean = 15·17
--- Lower Third Mean = 4·67

CORRECT RESPONSE SCORE : Acceptable Word Method

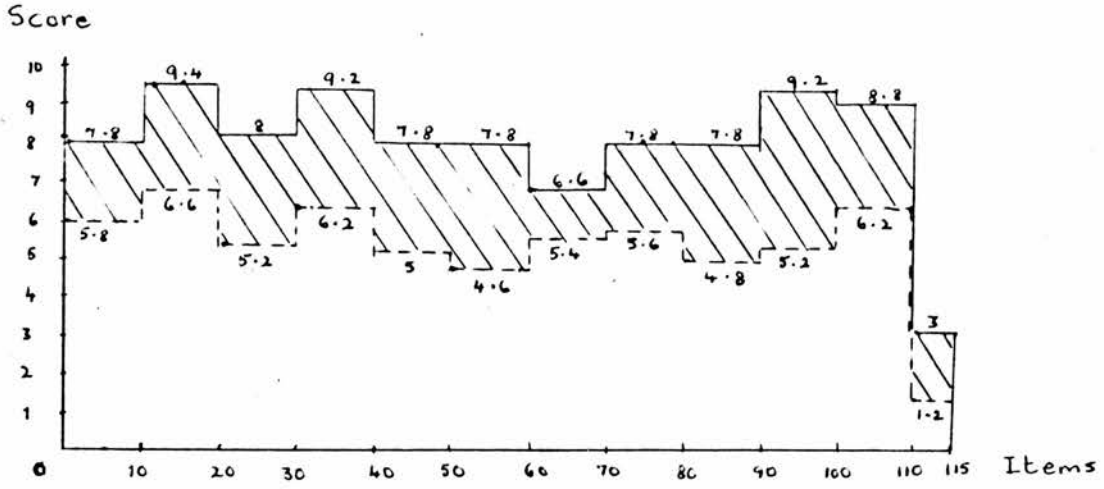


Fig. 22 PASSAGE 4A N = 16 —— Upper Third Mean = 93·2
--- Lower Third Mean = 61·8

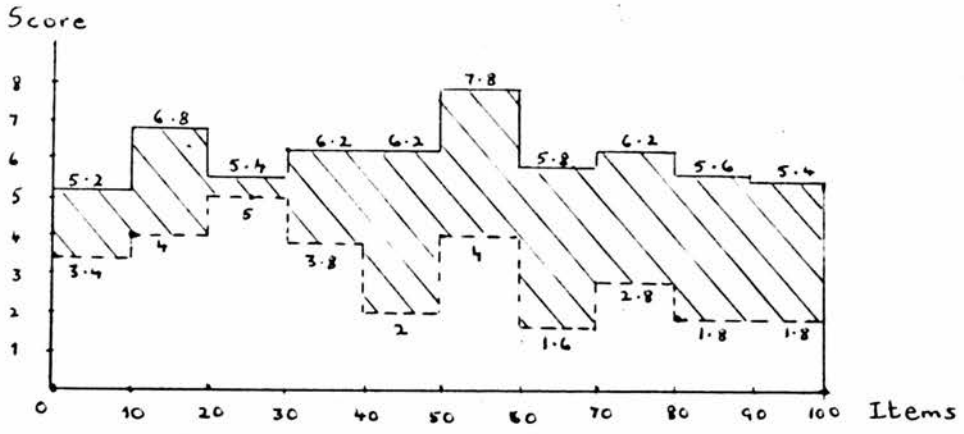


Fig. 23 PASSAGE 5A N = 16 —— Upper Third Mean = 60·6
---Lower Third Mean = 30·2

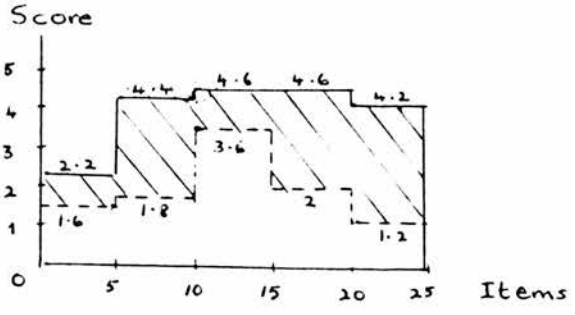


Fig. 24 PASSAGE 4B N = 16

— Upper Third Mean = 20

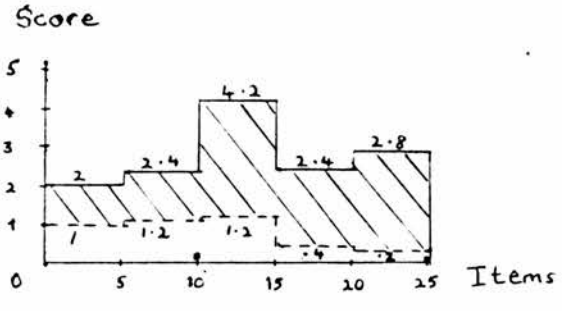


Fig. 25 PASSAGE 5B N = 16

Upper Third Mean = 13·8

INCORRECT RESPONSE SCORE

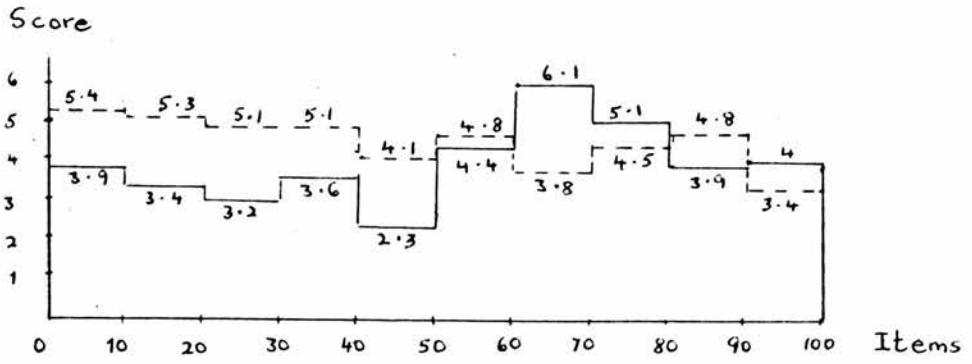


Fig. 26 PASSAGE 2A N = 49 —— Upper Third Score = 39.9
--- Lower Third Score = 46.3

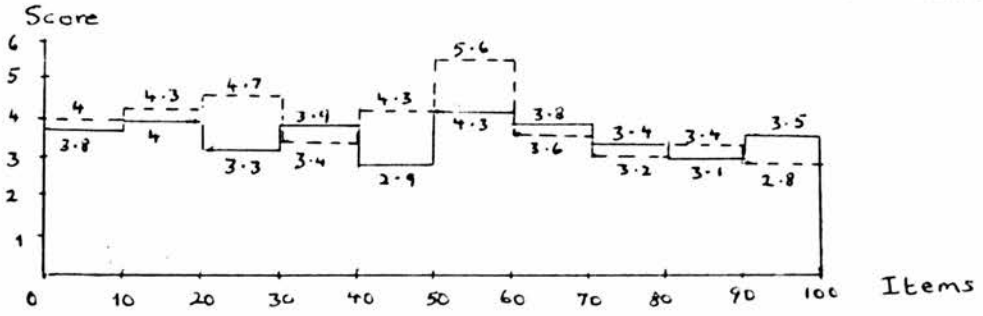


Fig. 27 PASSAGE 3A N = 36 —— Upper Third Score = 36.08
---- Lower Third Score = 39.17

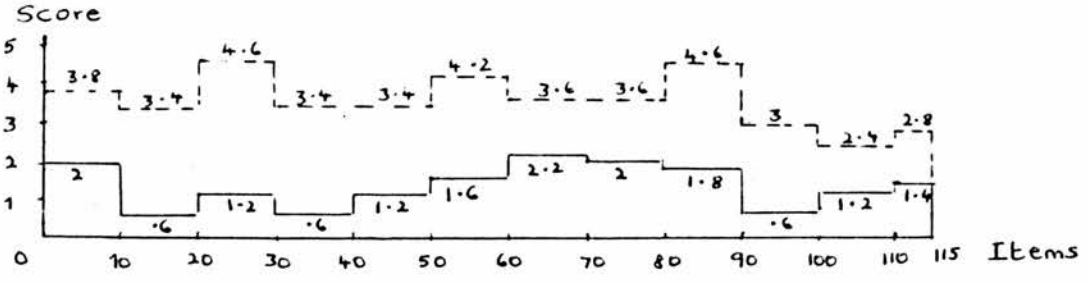


Fig. 28 PASSAGE 4A N = 16 —— Upper Third Score = 16.4
---- Lower Third Score = 42.9

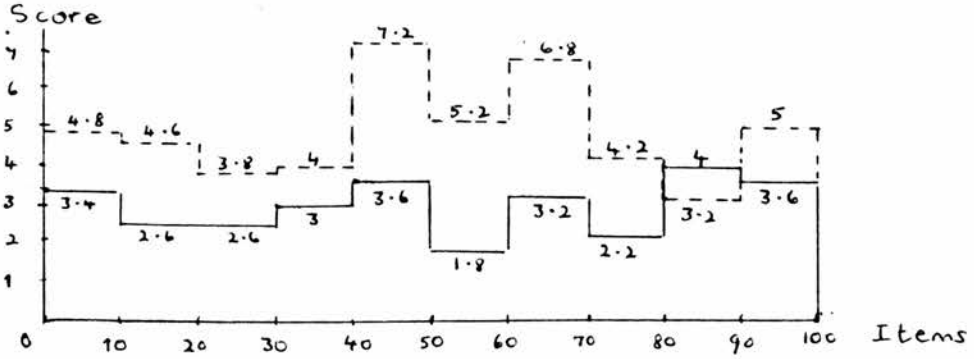


Fig. 29 PASSAGE 5A N = 16 —— Upper Third Score = 30
---- Lower Third Score = 48.8

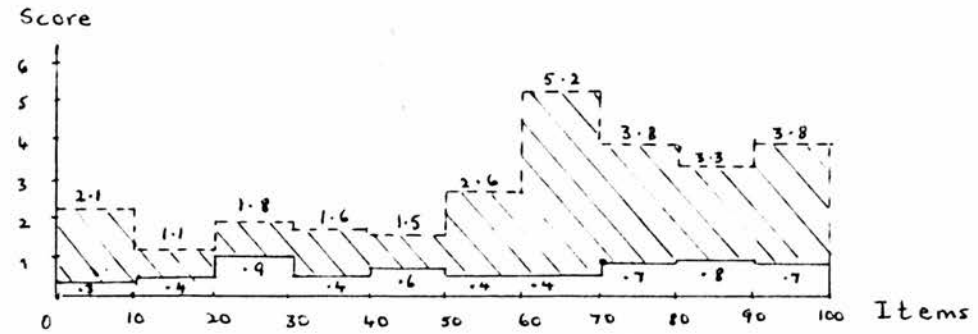


Fig. 30 PASSAGE 2A N = 49 —— Upper Third Mean = 5·25
--- Lower Third Mean = 26·80

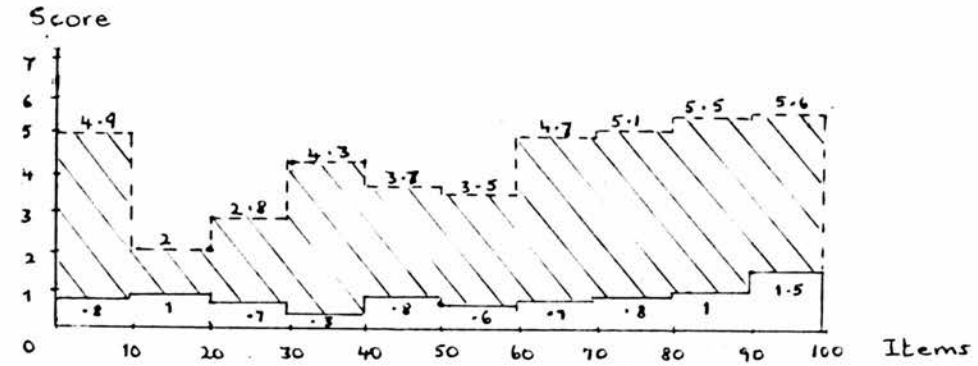


Fig. 31 PASSAGE 3A N = 36 —— Upper Third Mean = 8
---- Lower Third Mean = 42·08

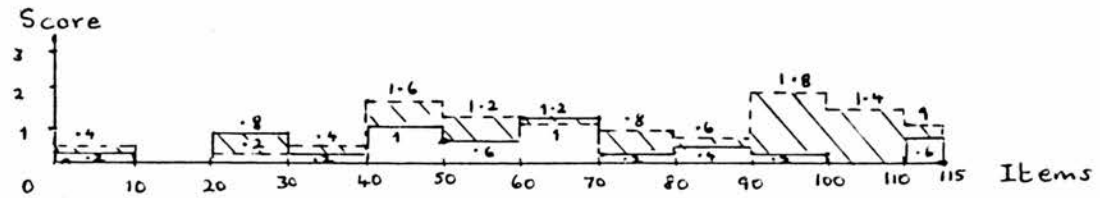


Fig. 32 PASSAGE 4A N = 16 —— Upper Third Mean = 5·4
---- Lower Third Mean = 10·4

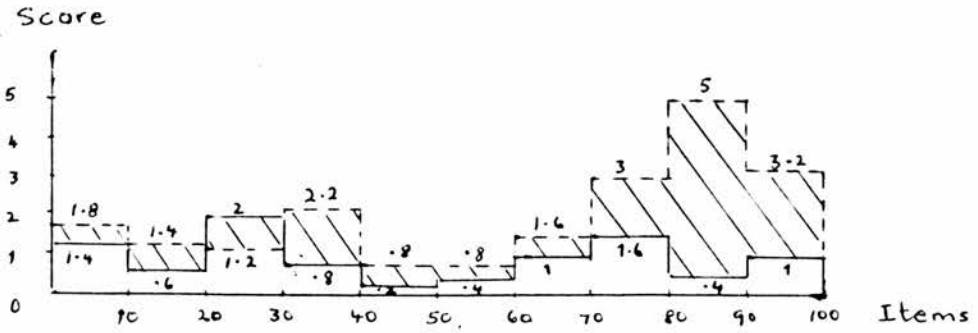


Fig. 33 PASSAGE 5A N = 16 —— Upper Third Mean = 9·4
---- Lower Third Mean = 21

INCORRECT RESPONSE SCORE

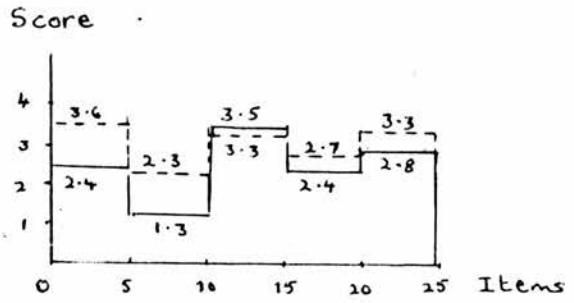


Fig. 34 PASSAGE 2B N = 49

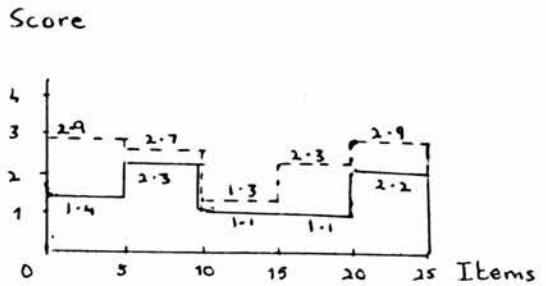


Fig. 35 PASSAGE 3B N = 36

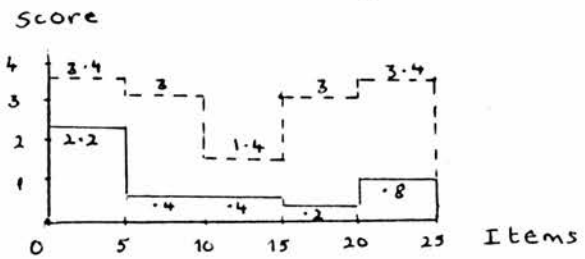


Fig. 36 PASSAGE 4B N = 16

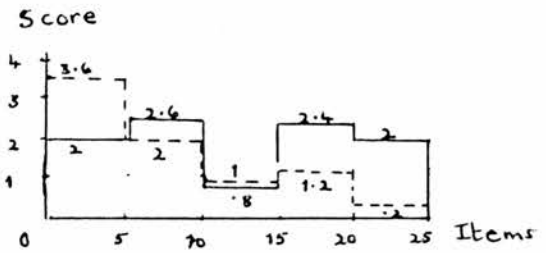


Fig. 37 PASSAGE 5B N = 16

OMISSION SCORE

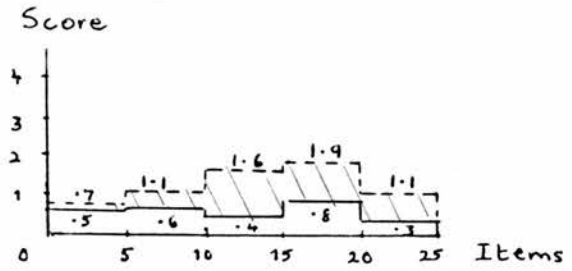


Fig. 38 PASSAGE 2B N = 49

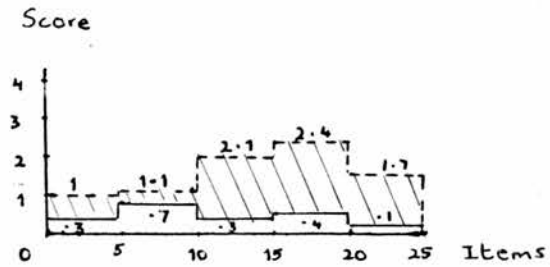


Fig. 39 PASSAGE 3B N = 36

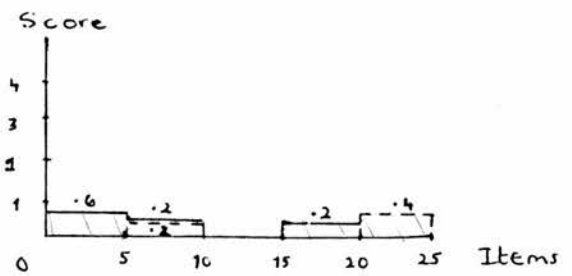


Fig. 40 PASSAGE 4B N = 16

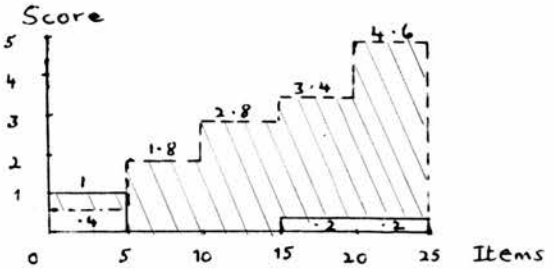


Fig. 41 PASSAGE 5B N = 16

Key : Upper Third Score ——
Lower Third Score - - -

8.3 SUMMARY

To conclude, this chapter presented the results of Experiments III and IV, the former concerned with the central issue of this thesis that process and product were different approaches to the investigation of reading ability, both providing complementary information. The latter experiment proved that length of text was a variable contributing to better discrimination between levels of proficiency when reading in a second language, and that longer texts were more reliable predictors of processing other long texts, an issue of particular relevance in our thesis and for our NNS population. Also evident was that texts graded for comparable difficulty when judged by readability measures were comprehended with varying levels of facility when measured by cloze tests.

The results of these and earlier experiments will be discussed in the next chapter.

SECTION FIVE

IMPLICATIONS OF THE INVESTIGATION

Chapter Nine	Discussion of Results and Development of the Reading Performance Profile.
Chapter Ten	The Model Reviewed
Chapter Eleven	Conclusion

CHAPTER NINE

DISCUSSION OF RESULTS AND DEVELOPMENT OF THE READING PERFORMANCE PROFILE

9. GENERAL REMARKS

In this chapter the results obtained from the experiments are first briefly discussed. This is followed by their application to the hypotheses investigated and the development of the reading performance profile based on the experimental evidence obtained. Implications for the reading model will be examined in the next chapter along with suggestions for further research.

Measurement of any kind and especially so in the educational field is considered as the elicitation or collection of data which is interpreted either in entirety or selectively in accordance with the tester's theoretical frame of reference. It is the latter that must, to a large extent, compensate for the inadequacies of the sampling of data and the various variables that one recognises are present, but over which there is little control. Although our techniques are crude and often inaccurate, they do attempt to reflect the processes we aim to test, and their validity depends to a large extent on their ability to stimulate the activities, processes and skills we propose to measure. The reason for measurement as we saw in Chapter Six dictated the choice of techniques that we selected.

The data we elicited reflected the reading ability of non-native speakers at an advanced level; and the adequacy of the measures was justified by the frame of reference set up through the reading model acting as a theoretical construct of what we hypothesised occurred during the activity. But before we could use the explanations to justify interpretation of the data it was necessary to obtain experimental validation of the measures used. 'If the reliability of a test ensures its consistency, validity ensures its meaningfulness' Davies (1977:58). Davies goes on to list five kinds of validity, face, predictive, concurrent, content and construct. Ingram combines predictive and concurrent under the general term 'pragmatic' (1977:25). The reading model was used to provide construct validity, while content validity was established by analysis and discussion of techniques and materials used in the experimental tasks. We suggest that the choice of long texts provided higher face validity than the three or four short passages often used in proficiency batteries.

What we present here is evidence of pragmatic validity based on reliability estimates and using the ELBA and EPTB as criterion measures of external validity. Each experiment is discussed independently at first.

9.1 EXPERIMENT 1 : PROCESS VERSUS PRODUCT - INFLUENCE OF ORDERING

The techniques used in this experiment were cloze and multiple-choice. In the first part, the difference in

means scored on a multiple-choice test given to two homogeneous samples was not significantly different when one group was given a cloze test based on the same passage prior to the multiple-choice test. So that although one of the groups was exposed to the text for a second time, the requirements of the two tasks compared in the experiment were so different as to have negligible direct transfer effects.

In the second part of this experiment the influence of a multiple-choice task on cloze was examined, and the mean was again not significantly different when compared with that obtained from the other half of the randomly grouped sample who had not done the multiple-choice test.

With satisfactory reliability and validity coefficients, these results produced confirmatory evidence that the two techniques were indeed measuring different aspects of the reading activity. This was inferred by the absence of significant effect due to ordering. Had the two techniques or tasks been measuring the same aspect, the effects of prior exposure to the test materials would have resulted in significant gains due to ordering.

9.2 EXPERIMENT II : PROCESS VERSUS PRODUCT - CLOZE VERSUS SUMMARY

Although the general area of investigation was similar to that in Experiment I, this second experiment was specifically designed to compare the results of techniques used to evaluate the process and the end-product

of reading. For the latter we used a summary task requiring paraphrase by recall of the contents of working memory. We postulated that results of processing during input were collated and organised in working memory by the subject using skills and strategies necessary for the task set.

The summary, generally accepted as a valid comprehension testing device, is however an open-ended task, and the results were inevitably affected by the subject's productive ability in the second language. Although the criteria for scoring were dictated more by comprehensibility than conformity to standard English, it was recognised that the subject possessing higher linguistic competence in English was better able to concentrate on communication of meaning instead of engaging in a memory search for relevant language forms necessary to convey the ideas in the second language.

Had the techniques used been regarded as alternative measures of the same ability, we would have had expectations of substantially high correlations when both referred to the same text. But if they were as we stated, two different aspects of the entire complex making up reading ability, then the shared variance would be considerably lower. The results supported our predictions. The correlations between cloze and summary when based on the same passage were higher than when obtained by using two different texts, although they were still far short of unity.

Correlations based on the same text were ·65 and

.75 (using exact and acceptable word methods). Correlations based on different texts were .58 and .64. But these results must be viewed with caution because of the relatively low reliability of one set of data (See Section 7.2.2.1). This is a disadvantage often found with open-ended tasks that have construct validity, but which are operationally prone to low statistical reliability because of the subjectivity accompanying the scoring procedures. Also supported was the hypothesis that contents of working memory are influenced by processing and comprehension of text during input. When the input was in the form of an unmutilated passage, the mean score for the summary task was higher than when the input was rendered more difficult by presenting it through a cloze task.

The mean for unmutilated input was 10.57 as compared to 7.73 via a cloze-type input. What is of interest here is however the correlation coefficients with the external criterion. Though samples were small, a correlation of .53 significant at .05 level was obtained with EPTB total scores. With ELBA total scores however, correlations were zero for the two small samples, although they approached a five percent level with the reading comprehension sub-test for both. One possible explanation could be the large number of items in ELBA that were aimed at testing aural skills, 150 out of 270. The comparable figure for EPTB was 96 out of 192 items. But the discrepancy is indeed disturbing between a test of receptive competence and a productive one, and there might be a case for including

something of the latter to make the entire proficiency battery more comprehensive, if similar results are obtained from further data.

But as we have already stated, the major weakness with this experiment lay in the lack of reliability estimates caused by the subjective nature of the method used to match the list of ideas contained in the passages used for the summary, with the responses produced by the subjects.

But this experiment did vindicate much of the criticism made against cloze as a measure of comprehension related more to the sentence level than to on-going discourse, and which we had earlier investigated in our informal experiment reported in Chapter Seven, Section 7.1.1. Had this charge been true, no summary would have been possible. But we do not claim that cloze is a good measure of 'reading comprehension' as the latter term is normally used, since it generally refers to understanding of the total completed input. For this, tasks like the summary, multiple-choice or open-ended questions or paraphrase etc. are more suited. Alternatively if the cloze technique was selected, it would be a good test of reading comprehension if it came after the reading and was based on a paraphrase of the text. With a time-limit imposed on the cloze task as we used it, the understanding we tested was primarily understanding at first exposure <u>during</u> the reading intake. There was little time for re-assessment of the overall text although it is true that, as with all timed tests, those who were already

proficient in the language and efficient in reading ability probably had more time for cognitive processes associated with the text as a whole than those whose linguistic ability was less competent. We will however have something more to say about speed of performance in the next chapter.

9.3 EXPERIMENT III : PROCESS VERSUS PRODUCT - CLOZE VERSUS MULTIPLE-CHOICE

9.3.1 The Cloze Test

This was the main experiment in which we obtained impressive testimony regarding both the reliability and validity of the cloze test as a measure of reading ability during the process stage.

The reliability coefficient for the British sample was .90, .96, based on verbatim and acceptable word methods, and is similar to those reported for large scale tests.

Of interest is the similarity of the validity coefficients obtained here as compared to results of cloze experiments with other standardised tests of proficiency in English as a second language and briefly described in Section 3.2. Darnell (1968) reported a correlation of .84 between a total of 200 items based on 4 cloze tests - scored by the exact word method - and the total score of TOEFL, with a group of 48 non-native speakers. Oller (1971) obtained a correlation of .75 and .83 between a cloze test - scored by exact and acceptable word methods - and the UCLA, ESL Placement Examination, with 35 non-native speakers. We obtained a correlation of .84, .87, - scored by the two

systems - between our 100-item cloze test and total scores on ELBA, with a sample of 45 non-native speakers. With the EPTB we obtained a correlation of ·66, ·72 for a sample of 51 NNS. This is identical with results on some preliminary work reported by Alderson (1976)[1] who also obtained a correlation of ·66 between his cloze tests and EPTB. All correlations were significant at ·001 level.

Analysis of responses revealed significant differences between the upper third on the one hand, and middle and lower thirds on the other. When scores based on the two scoring systems were compared it was evident that the upper third made substantial gains by the more relaxed method of scoring, whereas the increase in mean score was far less marked with the less proficient reader. Accompanying the gains was a lower correlation between verbatim and acceptable scores for the more proficient reader.

Incorrect responses did not reveal any marked trend apart from differences in quantity of wrong responses. There was a slight fall in incorrect responses with the lower group reflecting a slackening of effort, towards the latter half of the test.

By contrast the omission scores revealed a distinctly different patterning of performance with length having little or no effect on the more proficient readers, but having a very noticeable influence on the lower third readers. The definite increase in number of omitted items

---

1. Doctoral candidate at Edinburgh University.

after the 60 to 70 item block (which corresponded to the 650 to 750 word mark) could well reflect the breaking down of communication between writer and reader. It was possible that this length coincided with some kind of 'attention span' in reading at this level.[1] With the more proficient readers a lapse in attention would be followed by renewed effort, but this might well prove more difficult with those who were less linguistically competent. The reader who lacked exposure and training in English had probably not evolved the skills and strategies necessary for reading texts of this length with ease and efficiency.

As the results indicated the analysis of performance could be done in a number of ways. Along with the usual item analysis procedures, items that discriminated could be classified by the word class of the item deleted. When total correct scores of the British sample were classified into functional and content, it was found that the distribution of 8345 responses was as follows:

Correct response for <u>functional</u> items: Score = 5346.
Mean = 28·14 for 51 items, i.e. a success score of 55·17%.
Correct response for <u>content</u> items: Score = 2999.
Mean = 15·78 for 49 items, i.e. a success score of only 32.21%.

Earlier studies by Oller (1973:107) with NS had reported that functional items were easier to complete. This was also evidently true of NNS at this level.

---

1. Suggestion made by Dr. Alan Davies in personal communication.

Another alternative approach was by comparing the number and type of items that discriminated in the first part of the text with those that discriminated in the latter part. The following outline reveals that with length the number of items discriminating increased, and although functional items are easier on the whole, less proficient readers failed to respond successfully to 15 out of 17 functional items towards the end of the text.

First 30 items	
Items deleted:	Items discriminating:
15F + 15C = 30	8F + 4C = 12 i.e. 40%
Next 40 items	
Items deleted:	Items discriminating:
19F + 21C = 40	15F + 13C = 28 i.e. 70%
Last 30 items	
Items deleted:	Items discriminating:
17F + 13C = 30	15F + 8C = 23 i.e. 76·67%

TABLE 59 : GENERAL ANALYSIS OF CLOZE RESULTS (British Sample)

F = Function C = Content

If reduction in performance output was solely a matter of speed then the first seventy items or even the first thirty would have revealed comparable results. But as our analysis showed the test proved to be a good discriminator between subjects who obviously possessed different levels of proficiency in reading the second language. This point is illustrated very clearly by the following analysis of performance in Table 60.

UPPER THIRD N = 63			
	First 30 items	Middle 40 items	Last 30 items
Correct	1253 $\bar{x}$ = 19·89	1628 $\bar{x}$ = 25·84	1275 $\bar{x}$ = 20·24
Incorrect	536 $\bar{x}$ = 8·51	754 $\bar{x}$ = 11·97	474 $\bar{x}$ = 7·52
Omitted	101 $\bar{x}$ = 1·60	138 $\bar{x}$ = 2·19	141 $\bar{x}$ = 2·24
	1890	2520	1890
MIDDLE THIRD N = 64			
	First 30 items	Middle 40 items	Last 30 items
Correct	815 $\bar{x}$ = 12·73	1107 $\bar{x}$ = 17·30	768 $\bar{x}$ = 12
Incorrect	806 $\bar{x}$ = 12·59	940 $\bar{x}$ = 14·69	657 $\bar{x}$ = 10·27
Omitted	299 $\bar{x}$ = 4·67	513 $\bar{x}$ = 8·02	495 $\bar{x}$ = 7·73
	1920	2560	1920
LOWER THIRD N = 63			
	First 30 items	Middle 40 items	Last 30 items
Correct	571 $\bar{x}$ = 9·06	603 $\bar{x}$ = 9·57	325 $\bar{x}$ = 5·16
Incorrect	852 $\bar{x}$ = 13·52	1079 $\bar{x}$ = 17·13	482 $\bar{x}$ = 7·65
Omitted	467 $\bar{x}$ = 7·41	838 $\bar{x}$ = 13·30	1083 $\bar{x}$ = 17·19
	1890	2520	1890

TABLE 60 : ANALYSIS OF RESPONSES FOR UPPER, MIDDLE AND LOWER THIRDS : CLOZE TEST P1 N = 190

Of interest is the fact that in the first part of the test, the middle and lower third groups had roughly the same number of incorrect items. By the middle of the passage, there was a marked increase in incorrect and omitted items for the lower third group. By the end of the test, we see that the three groups have become definably distinct from each other in performance.

What these alternative approaches to analysis of data have shown is the value of this technique as both a

measurement and an investigative device. We have been able to compare performance not only between levels of ability, but more importantly, within levels. This latter approach enabled us to compare variations in response from one part of the text to another.

9.3.2 The Multiple-Choice Test

The multiple-choice test presented problems often associated with the technique. As our results showed, Version 1 had a reliability of ·75, and a validity coefficient of ·66 with total ELBA, the latter significant at more than ·05 level. Version 2 which was developed as a result of the earlier try-out had a lower reliability of ·65 and a validity of ·5 with the criterion measure EPTB. This was significant at more than ·05 level too, but again this is what one would expect between a total proficiency test and a measure of one complex skill. Seventeen of the twenty items were reasonably satisfactory. Item 8 which involved a mathematical calculation continued to attract the majority to the wrong distractor and had to be replaced in Version 3.

Items 13 and 15 were unsatisfactory from the beginning. The wording of the questions proved difficult and in spite of alterations made to improve the quality, the tester remained dissatisfied. They were retained because they aimed at evaluating particular skills referred to in the taxonomy. Item 13 sought to determine whether subjects were capable of identifying if a 'factual' statement was based on inference or on positive evidence. There were

many instances in the text where the author used phrases like 'may be said to be', 'seems to have begun', 'we can guess' '...was probably...', andit seemed pertinent that NNS recognise fact from inference. The difficulty experienced with Item 15 was probably due to the way in which the tester listed the separate factors necessary for the origin of the state. The fault here lay in the fact that each factor was not clearly separate but provided an overlap which may have confused the subject. The aim in this question was similar to that of Item 4. Features to be included in the definition of a 'state' had to be identified.

The results of these versions were presented in terms of the taxonomy since this was the basis on which the questions were prepared.

Version 3 was mainly exploratory and represented an area of special interest to the author. This was concerned with text structure or organisation of text, subsumed under "other skills" (Section 6.2.3.2.). In the summary experiment (Section 7.2.2.), the list of ideas for each of the two passages used, was prepared by an analysis of the text in terms of its discourse function. This was done by assigning to each sentence, or part of a sentence its function under labels such as: proposition, evidence, summing-up, qualification etc. Appendix 3 discusses this in greater detail. In most studies where 'text structure' has been identified as one of the areas included for testing, Davis (1944 and 1968) and Barrett

(1968),[1] this term has referred mainly to chronological sequence of events. Item 19 of the two earlier versions was of this type. It was decided to explore this further in Version 3 for which seven new items were prepared five concerned with text structure. Items 21 to 25 were added and Item 8 replaced. Because one more item was to be tried, Item 6 (a good item) was replaced. No additional time was allowed which was probably a mistake. As the results indicated, of the seven items tried, only three worked with the small sample used. This however does appear to be an area that I think might well prove worth investigating. How a text is organised and what contributes to cohesion may not be self-evident to the NNS although it is likely that this aspect is more suited to teaching than to testing. As reported earlier this version had .60 reliability but apart from concurrent validity with cloze, remained untested for pragmatic validity with other criterion measures. This was an experimental version and did not affect the results of Experiment III. But before moving on to a discussion of the correlations between the results of the cloze and multiple-choice tests, I would like here to briefly present the results of all three versions on the basis of questions aimed at the ability to abstract information based on detail. The overall success in these items is compared with the ability to abstract information relating to more generalised concepts like summary, recognising fact from inference, re-arranging events in chronological order etc.

---

1. Quoted by T. Clymer (1968).

In Version 1 and 2, there were 12 questions on detail, and 8 on general concepts and the scores were distributed as follows:

Version 1 : 20 questions N = 60

12 questions on detail : Score = 481
Mean = 8·02 or 66·81%

8 questions on generalised concepts : 281
Mean = 4·68 or 58·54%

Version 2 : 20 questions N = 152

12 questions on detail : Score = 1030
Mean = 6·78 or 56·47%

8 questions on generalised concepts : 618
Mean = 4·07 or 50·82%

Version 3 : 25 questions N = 22

12 questions on detail : Score = 150
Mean = 6·82 or 56·82%

13 questions on generalised concepts : Score 145
Mean = 6·59 or 50·70%

Of interest is the score obtained on Items 6, 19, 21, 22 23, 25, in Version 3, which were concerned with the general organisation of the text and so indirectly of the author's organisation of the content. The score obtained was 47, giving a mean of 2·14 or 35·61% success. Because the items failed to discriminate or even function adequately, the results of this part are only to be taken as representing an area that needs further investigation. This might be done by manipulation of various linguistic devices open to the writer in English.

Since in all three versions of the multiple-choice test, a record of omissions was made, a brief analysis is

presented in Table 61. The omission scores are presented in terms of the taxonomy and it is interesting to see how they parallel results obtained from the correct score working in inverse ratio to level of ease.

TAXONOMY	VERSION 1 N = 60		VERSION 2 N = 152		VERSION 3 N = 22	
		score		score		score
Localised distractors	(7 items)	14	(7 items)	51	(9 items)	13
vs						
Non-localised distractors	(13 items)	33	(13 items)	164	(16 items)	36
Factual	(3 items)	3	(3 items)	22	(2 items)	1
vs						
Referential	(4 items)	5	(4 items)	17	(4 items)	2
vs						
Inferential	(7 items)	20	(7 items)	82	(7 items)	16
vs						
Evaluative	(6 items)	19	(6 items)	94	(12 items)	30
Detail	(7 items)	22	(7 items)	87	(12 items)	16
vs						
Generalized Concepts	(13 items)	25	(13 items)	128	(13 items)	33

TABLE 61 : ANALYSIS OF OMISSIONS IN THE MULTIPLE-CHOICE TESTS

9.3.3 Cloze Versus Multiple-Choice

The results of the third part of this experiment compared the performance of subjects on the two tests used to measure the process and product of reading. The concurrent validity coefficients that emerged revealed a large amount of unshared variance. These could be satisfactorily explained by reference to our model, and were used to provide

support for the hypothesis that the two techniques used were indeed evaluating different aspects of the ability being investigated.

9.4 EXPERIMENT IV : INVESTIGATION OF LENGTH AS VARIABLE

All the eight texts used in this experiment had been judged by readability formulae as having a level of difficulty appropriate to college level reading except for P2A, P2B and P4A which by FRY and SMOG formulae were rated as easier. The cloze results however placed Passage 2A and 2B at the more difficult end of the scale for the sample who sat these tests.

Since different samples sat the tests the only comparison we could make that might be meaningful was by using the performance of these four samples on the P1 Cloze Test. Using this to rank the four groups we found that mean scores ranked the four long texts in identical fashion. Each group is identified by its test passage number.

GROUP	PASSAGE 1	RANK	LONG TEXT:A	RANK	SHORT TEXT:B	RANK
2	$\bar{x}$ = 36.71	3	$\bar{x}$ = 40.49	3	$\bar{x}$ = 6.73	4
3	$\bar{x}$ = 34.28	4	$\bar{x}$ = 37.53	4	$\bar{x}$ = 10	2
4	$\bar{x}$ = 66.29	1	$\bar{x}$ = 75.44	1	$\bar{x}$ = 15.63	1
5	$\bar{x}$ = 46.13	2	$\bar{x}$ = 53.44	2	$\bar{x}$ = 9.25	3

TABLE 62 : TEXTS GRADED BY CLOZE FOR DIFFICULTY

(1 Easy to 4 Difficult)

To permit a comparison between predicted difficulty levels of the various texts used (See Section 6.2.5.2.) and difficulty levels as they emerged on the basis of performance in cloze tests based on the passages, we present in Table 63 the combined results of Table 7 (Section 6.2.5.2.) and Table 62. Texts are ranged from easy to difficult.

PREDICTED RANKINGS OF LONG TEXTS			OBTAINED RANKINGS
Readability Formulae	Deletion Count	Lexical Variation	Based on cloze scores
2A	1	2A	4A
1	2A	4A	5A
4A	4A	1	1
3A & 5A	5A	3A & 5A	2A
-	3A	-	3A
SHORT TEXTS			
2B	2B	5B	4B
4B	3B	4B	3B
5B	5B	3B	5B
3B	4B	2B	2B

TABLE 63 : GRADATION OF TEXTS (from easy to difficult)

As is evident, apart from the P 3A which was predicted difficult and which turned out to be so, none of the three measures used could be termed reliable. None of the readability formulae - including Bormuth's cloze formula - were successful in predicting outcomes. Neither did deletion count, or the lexical density and lexical variation counts have much in common with the rankings as they emerged. This highlights the basic difference between the measures.

Whereas readability formulae are predictive measures determining the level of the text, cloze is a measurement device that determines the ability of the population sample who read the text. It gives no clue regarding the difficulty of materials except in terms of the population sample. The fact that the same text P1 provided so many different cloze results indicated that the population taking the cloze tests was indeed spread from proficient at one end of the scale, to NNS at the other end who found the text too difficult, and were therefore lacking in the skills necessary for successful reading at this level.

Reliability for the long texts used in this fourth experiment ranged from ·75 to ·90 by exact word standards, and from ·89 to ·94 by acceptable word criteria. However, not unexpectedly with only 25 items deleted, the shorter tests displayed a much lower range of reliability figures, ranging from ·41 to ·73 by exact word methods and from ·39 to ·82 by acceptable word criteria.

The results showed that when responses were scored for correct, incorrect and omitted items, the differences between levels of proficiency were most clearly indicated by the omitted items when the text was not too easy, that is, with all texts except 4A and 4B. Passage 4A which had the highest mean score for correct responses was also the longest text used, with 115 items to be filled in the same time as provided for the other texts with only 100 items. What is interesting is that the rate of omission was negligible demonstrating very clearly that given

sufficient ability, the time of 45 minutes was adequate for the task; and that the rate of omission was more a reflection of communication breakdown caused by insufficient reading ability to process the text adequately, than lack of sufficient time to complete the task.

As we had already reported, all the evidence obtained here indicated that if students were to be required to read long texts as part of academic work, a test using a longer text had better predictive validity than texts of only 250 to 300 words or less, such as are used in tests of proficiency.

Additional data is provided by comparing the results of individual performance of a small sample of subjects in Table 64. In each block of ten items, the central mark indicates the correct score by acceptable word standards. To the left we have the number of incorrect items, and to the right, the number of omitted items. Attention is drawn to performance both at the beginning and in the second half of the text. Had the test been only 150 or 200 words long with only ten or twenty items deleted, it would have been difficult to differentiate between levels of proficiency in many cases. With length the task was easier.

No.16 (UPPER THIRD)	7 3	5 4 1	4 6	4 6	4 6	4 6	6 3 1	5 5	5 4 1	4 6	48 3 = 49
No.12(MIDDLE THIRD)	8 2	7 3	5 5	5 5	4 6	5 5	8 1 1	9 1	5 2 3	5 1 4	61 4 = 35
No.14 (LOWER THIRD)	6 4	6 4	8 1 1	3 7	3 7	7 3	3 6 1	7 3	9 1 0	6 2 2	58 10 = 32

TABLE 64 : PASSAGE 2A : INDIVIDUAL SCORES

All the three subjects sat this long text as the second of the pair, having done the shorter text first; and because of the similarity in topic, had benefited by the reading of the earlier short passage. The numbers identify the subjects as recorded in Appendix 4.D.4.1 page 194. See page 199 of the Appendices for an overall record of the subjects reported in Table 65.

No.11 (UPPER THIRD)	3	6	5	4	3	7	5	7	4	7	49 =	51
No.15 (LOWER THIRD)	5	4	5	6	2	4	1	2	1	1	46 =	31

TABLE 65 : PASSAGE 5A : INDIVIDUAL SCORES

Both subjects did P5A as the first of the pair, with 5B following. Further data is included in Appendix 4.E.2. We can see that correct and omitted scores reveal distinct trends, incorrect scores being less informative at least in terms of quantitative data.

Let us turn now to the interpretation of the results with reference to the hypotheses investigated. For this purpose the evidence will be viewed collectively rather than piecemeal regardless of the particular experiment that provides the necessary data.

9.5 EVALUATION OF HYPOTHESES

Hypothesis 1. 'Tests of cloze requiring the successful completion of a large number of deleted items will correlate substantially with standardized tests of English as a second language, thereby establishing the degree of relationship between reading ability and linguistic proficiency in NNS.'

The high reliability and validity coefficient obtained in the cloze test (Passage 1) with 100 deleted items confirmed the hypothesis that cloze as a search process, activated stores similar to those evaluated in tests of proficiency that included sections on vocabulary and grammar etc. The implication was that cloze which had evaluated the reading process during input was a reliable and valid measure of the subject's linguistic proficiency. And to the extent that the latter affected ability to process verbal graphic input with ease, it reflected and provided a measure of reading ability.

Hypothesis 2. 'When cloze items are scored by verbatim i.e. exact word replacement and by acceptable word methods, the gains in scores will be substantially higher for the more proficient readers. Differences in scoring systems can alternatively be measured by correlation, with lower values obtained for the better readers as compared with the less proficient readers.' Although Oller (1973:109) had reported that there was something counter-intuitive about scoring the performance of NNS by the verbatim method alone, we found that correlations between the two systems were so high as to suggest that the less tedious exact word method was perfectly adequate for group results. This high correlation was obtained not just from the cloze test of the main experiment, but from all the other eight cloze tests used in the experiments. The range of coefficients lay between .94 to .99 for the five long texts, and from .89 to .97 for the four shorter passages.

These overall group correlations however blurred differences that emerged when we analysed the results in terms of upper, middle and lower groups to test the specific hypothesis. In the main experiment, the increase in mean score from lower to upper thirds was accompanied by a fall in correlation coefficient as hypothesised both for the British and Nuremberg samples. The implication here was clear. For the more proficient reader, the acceptable word system had much to recommend it if individual scores were important. It also yielded better reliability and validity coefficients. If, however, the reader was a non-native speaker of little proficiency, the setting up of the more elaborate scoring system was unnecessary.

Since the hypothesis referred to scoring techniques used for cloze, we used the remaining eight texts to re-test the hypothesis with additional data. Whereas the main part of the hypothesis was confirmed because the more proficient readers did substantially increase scores in all tests, this gain was not always accompanied by correlation coefficients in inverse ratio. The following details make the point clear. Since the eight texts apart from P1 were used with smaller samples, details for them are given only for upper and lower third groups.

TEST	GROUP	EXACT MEAN	ACCEPTABLE MEAN	GAIN	CORRELATION
P1	UPPER THIRD	48.25	65.97	17.72	·86*
	MIDDLE THIRD	31.7	42	10.3	·88*
	LOWER THIRD	19.19	23.79	4.60	·95*
P2A	UPPER THIRD	42.56	54.81	12.25	·92*
	LOWER THIRD	22.56	26.94	4.38	·86*
2B	UPPER THIRD	7.13	9.94	2.81	·27
	LOWER THIRD	3.38	3.63	.25	·91*
P3A	UPPER THIRD	42.75	55.91	13.16	·95*
	LOWER THIRD	14.83	18.75	3.92	·95*
3B	UPPER THIRD	12.33	15.17	2.84	·82*
	LOWER THIRD	3.83	4.67	.84	·94*
P4A	UPPER THIRD	66.6	93.2	26.6	·95*
	LOWER THIRD	40.4	61.8	21.4	·87*
4B	UPPER THIRD	11.4	20	8.6	·76
	LOWER THIRD	4.4	10.2	5.8	·66
P5A	UPPER THIRD	39	60.6	21.6	·85*
	LOWER THIRD	21.8	30.2	8.4	·93*
5B	UPPER THIRD	9.4	13.8	4.4	·73
	LOWER THIRD	2.8	4	1.2	·83*

TABLE 66 : ANALYSIS BY TWO SCORING SYSTEMS FOR 9 CLOZE TESTS

The average gains as we see were always substantially greater for the more proficient group, ranging from 12 to 27 for the long texts and from 3 to around 8 for the short texts for the upper third group. The lower third group showed gains ranging from 4 to 21 in the long texts and from ·25 to around 6 in the shorter ones.

But a correlation between the two systems was not an alternative method of calculating gain since a fall in correlation coefficient did not always occur in inverse ratio to the gain. Our evidence confirmed that where group results were needed, exact scores were adequate as were results for non-proficient readers, but with proficient readers and an interest in individual results the acceptable word system was better. As is evident there was more to the hypothesis than appeared at first glance and the advantage of an investigative study lay in precisely this kind of freedom to explore the data more fully than would otherwise have been the case. Hypotheses 3 and 4 were both related specifically to length.

Hypothesis 3. 'The longer the text, the easier it is to discriminate between different levels of reading ability i.e. between more and less proficient readers.'

Hypothesis 4. 'The number of omissions representing failure to retrieve any word that might conceivably fit the constraints imposed by context, is likely to increase with length of text if the reader is not proficient in the language. This is likely to be accompanied by a reduction in speed of performance.'

As our results showed both hypotheses were confirmed and alternative ways of analysing the data showed this was borne out in various ways. By analysing the total response patterns, the graphs that accompanied all nine texts showed that one of the best indices was the omission score except when the text was too easy for the sample. With correct

word scores, the increase in length did increase the difference between levels of performance for all texts except P2A. This text which we had labelled inefficient and uneconomic because of the small number of items discriminating between good and poor readers, did however very successfully identify the two groups from the start. The difference did not increase with length, but it remained clearly defined throughout the text (Section 8.2.5. Fig. 18).

An alternative analysis looked at the number of items discriminating between the first half of the long text as compared with the second half, and the following data shows clearly that except for P2A, the other long texts all showed growing differences with length, with more items discriminating in the second half.

TEXT	ITEMS DISCRIMINATING IN FIRST HALF	ITEMS DISCRIMINATING IN SECOND HALF	TOTAL
P1	26	38	= 64
P2A	19	15	= 34
P3A	25	31	= 56
P4A	19	25	= 44
P5A	20	27	= 47

TABLE 67 : COMPARISON OF ITEMS DISCRIMINATING IN THE FIRST AND SECOND HALF OF LONG TEXTS

The increase in omitted items for all four texts (not including P4A) came after the same 600 to 750 word mark reported earlier in our discussion of results for P1 (Section 9.3.1) and supported the earlier suggestion that

perhaps this coincided with what might turn out to be the limit of an 'attention span' in reading at this level.

The implication here is that as we reported earlier, the longer text is a better predictor of performance for other long texts. And if there are marked differences in performance beyond the 700 word mark perhaps the reading comprehension sub-test in proficiency batteries ought to include two sections : One containing short texts for intensive study skills, and one including texts for extensive reading ability that would include this 'span'. The sample of individual scores (See Tables 64 & 65, and Appendix 4.B.2) showed how difficult it was to reliably separate levels of performance if the cloze test was short. To summarise, both hypotheses were confirmed and in addition we obtained ample evidence that justified our decision to use texts of 1000 words and more for investigating the reading ability of non-native speakers.

Hypothesis 5. 'Evaluation of the process via cloze though positively and significantly correlated with the evaluation of the product of reading, will have sufficient unexplained variance to confirm that they are reflecting two separate aspects of the reading activity.'

This aimed at validating one of the basic formulations of the model i.e. that process and product were two complementary, not alternative, approaches to the measurement of reading proficiency, each providing different evidence of ability and each capable of evaluation by techniques peculiarly suited to their separate tasks.

This assumption was supported by the following results which confirmed the hypothesis:

(i) absence of influence due to ordering.

(ii) the positive correlations between process and product which still left a large amount of unshared variance. Although the product stage was evaluated by two techniques, summary-writing and multiple-choice, correlations were never large enough to warrant the assumption that the tests were all doing the same job.

The use of the multiple-choice technique to evaluate the product of reading introduced greater statistical reliability and validity to the inferences made from the results, helping to confirm the hypothesis that the techniques selected had indeed activated related though identifiably different areas of the cognitive processing system.

Hypothesis 6. 'Questions of detail explicitly stated or implied will be easier for all readers than questions involving more generalised concepts.'

As the discussion of results in Section 9.3.2. revealed, this hypothesis was confirmed in all three versions of the multiple-choice test to which it related. The success score for each category is reported in terms of percent for ease of comparison.

Version 1 : 66.81% for detail versus 58.54% for generalised concepts

Version 2 : 56.47% for detail versus 50.82% for generalised concepts

Version 3 : 56.82% for detail versus 50.70% for generalised concepts

Studies that had attempted to factor analyse results obtained on the basis of similar taxonomies such as ours, had reported that no clear evidence emerged except for perhaps two areas, one related to vocabulary or word meanings, and the second a general verbal factor. Lunzer (1975:8) reported that he and his colleagues 'did not find any statistical techniques that were successful in separating out the separate skills [they] had tried to build into the tests' and that all pointed to a single factor that might be termed 'comprehension'.

We made no attempt to analyse the tests in this way using only the mean score for our results. As we had stated earlier, the advantage of a taxonomy lay not in claiming that it represented independent sub-skills that together made up reading, an assembly model to use Guthrie's description (1973), but in viewing it as a taxonomy of objectives that represented various abilities and skills the subject needed to exercise if he was to succeed in abstracting information from the text read. The difference in mean scores confirmed that not only was detail easier but that of the skills involved in item solving, the recognition of factual and referential information proved much easier than questions that required inference and evaluation even with the more proficient readers forming the upper third group.

Hypothesis 7. 'Items with distractors taken from a localised context will be easier than items where the distractors occur in widely separate contexts.'

This last hypothesis had to be rejected on the basis of evidence obtained in all three versions of the multiple-choice test.

I would however like to draw attention to the confounding of two variables, one the location of distractors and the other, the use of the taxonomy in the preparation of items. It is possible that this hypothesis was not really tested as such. We should have compared the item difficulty based on location of distractors alone by having comparable items testing for factual information, inferential and evaluative abilities.

I would therefore suggest that theoretically the hypothesis retains its validity and force and failure to obtain supportive evidence here is due to inefficient testing. Equally obviously it has not been proven in the thesis.

Since the experimental evidence established the pragmatic validity of six of the hypotheses proposed to investigate the reading model for a non-native speaker of English, we used the results of performance in the Cloze test based on Passage 1 to see whetherit revealed characteristics that could be identified with proficient readers as opposed to those exhibiting evidence of lower reading ability in the second language. What we were looking for was a pattern of responses to the cloze test (which investigated the process of reading) that could be identified with a level of ability and which would represent what we designated as a 'reading performance profile'.

9.6 DEVELOPMENT OF THE READING PERFORMANCE PROFILE

In order to obtain a representative picture we based the profile on the mean scores obtained from the British sample that sat the cloze test. These scores reflected the rate of progress through the long passage and by analysis of correct, incorrect and omission scores would present an 'average' performance by non-native speakers. Our concern was to identify the proficient from the less able reader so that the latter could obtain remedial treatment if this should prove desirable. In order to determine what constituted 'adequate' reading proficiency we used the two measures ELBA and EPTB which had provided the necessary external validity for the test of cloze.

From the results obtained, the group of 45 subjects who had sat the ELBA, had the following mean score:

Cloze (Exact) = 37.49% (based on 100 items)

Cloze (Acceptable) = 51.11%

ELBA (Total) = 64.59% (based on 270 items)

From the 51 subjects who had sat the EPTB, the following mean score was obtained:

Cloze (Exact) = 27.82%

Cloze (Acceptable) = 36.69%

EPTB (Total) = 35.34 [standard score based on 192 items]

For the ELBA, a score of 70% is considered the minimum desirable for students wishing to attend courses in social sciences, while a score of 60% is suggested as perhaps adequate for students enrolling on physical science courses.

For the EPTB, a score of 40 is considered as the minimum desirable for all students.

Using these criteria as a rough guide we selected the upper third group who had obtained the following mean score to represent the norming group for a proficient reader. They had obtained an overall mean score of :

Cloze (Exact) = 48.25%

Cloze (Acceptable) = 65.97%

We inferred that this group possessed a level of reading proficiency sufficient for the demands imposed by an academic level course of instruction in English. Using the total score to identify the middle and lower third groups, their performance was taken to represent a level of performance inadequate for the requirements imposed by an academic level course conducted through the medium of the second language, English.

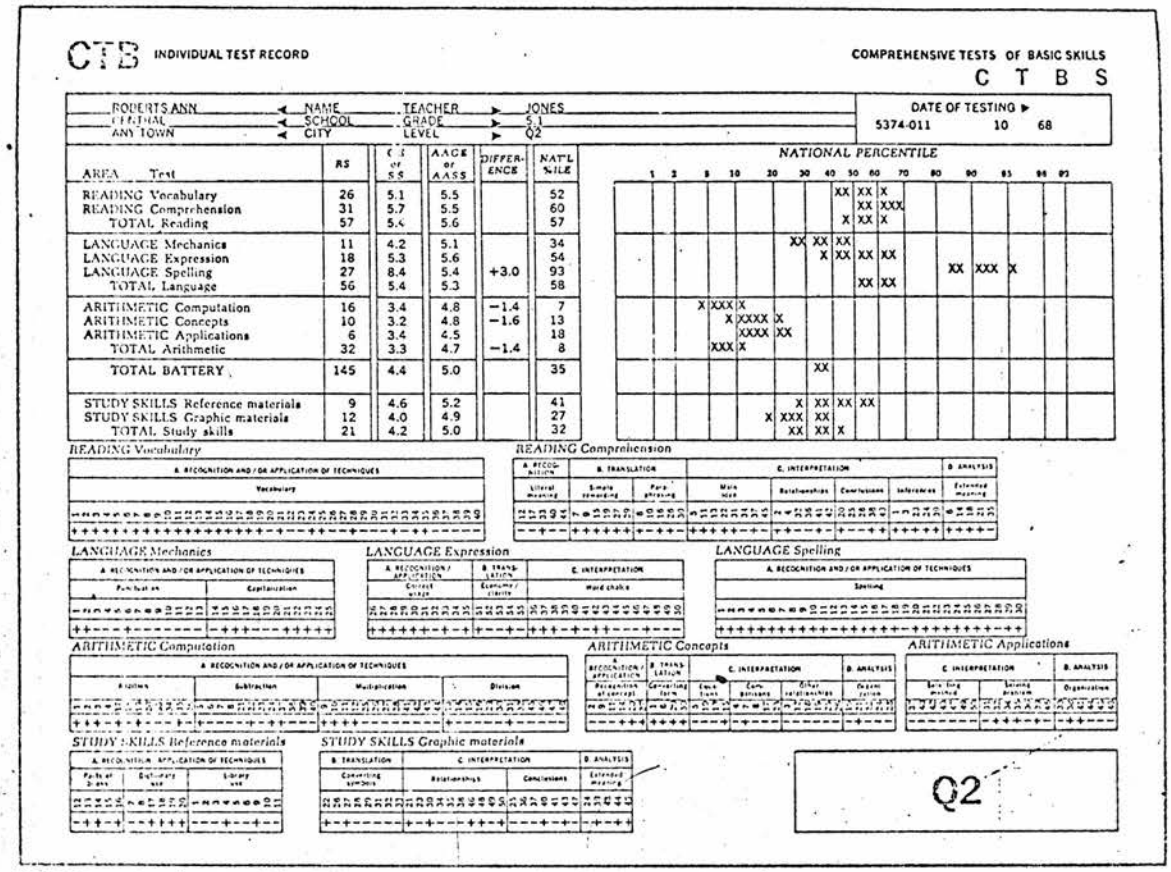
The concept of a profile applied to a pupil's ability is not new in educational circles, however it has generally been used to refer to the student's competence in different skills all of which together represent his performance. This is usually in various areas of study that according to some authority are designated as 'important' and therefore assume the role of desirable objectives. The record of learning is based either on a standardised battery of tests or on the teacher's rating. Sometimes the profile represents evidence of knowledge of content and sometimes of skills, and the profile enables the teacher to compare pupil performance against national or local norms. We

include two of these for information on the next page, the S.R.A. Pupil Profile and the CTB Individual Test Record.

Psychological studies also exist which purport to represent personality profiles. These are often based on questionnaire-cum-interviews.

What was proposed here however was something rather different. We focussed on one complex ability, reading; and by presenting a view of performance during the reading process, permitted observation of rate of progress in a long passage that experimental use had shown was a good discriminator between different levels of reading proficiency. Since the purpose of profiles is to permit evaluation of the individual in terms of some criterion that acts as a norm or criterion measure, what we presented was a cross-section of performance by three groups of readers who provided the norms for different levels of ability in reading. The correct score represented an accuracy score and the omission score represented speed in inverse ratio. Theoretically ability has to do with competence and subsumes comprehension of the text under investigation, so that we inferred the level of reading ability on the basis of test performance.

Having divided the group of 190 subjects of the British sample into three groups on the basis of total scores, their performance was analysed for correct scores (by acceptable word standards), incorrect scores and omitted scores. For the profile of performance, the test of 100



CTB INDIVIDUAL TEST RECORD

COMPREHENSIVE TESTS OF BASIC SKILLS

C T B S

ROBERTS ANN ◄ NAME TEACHER ► JONES
CENTRAL ◄ SCHOOL GRADE ► 5.1
ANY TOWN ◄ CITY LEVEL ► Q2

DATE OF TESTING ► 5374-011 10 68

AREA Test	RS	G.E. or S.S.	AAGE or AASS	DIFFERENCE	NAT'L %ILE
READING Vocabulary	26	5.1	5.5		52
READING Comprehension	31	5.7	5.5		60
TOTAL Reading	57	5.4	5.6		57
LANGUAGE Mechanics	11	4.2	5.1		34
LANGUAGE Expression	18	5.3	5.6		54
LANGUAGE Spelling	27	8.4	5.4	+3.0	93
TOTAL Language	56	5.4	5.3		58
ARITHMETIC Computation	16	3.4	4.8	−1.4	7
ARITHMETIC Concepts	10	3.2	4.8	−1.6	13
ARITHMETIC Applications	6	3.4	4.5		18
TOTAL Arithmetic	32	3.3	4.7	−1.4	8
TOTAL BATTERY	145	4.4	5.0		35
STUDY SKILLS Reference materials	9	4.6	5.2		41
STUDY SKILLS Graphic materials	12	4.0	4.9		27
TOTAL Study skills	21	4.2	5.0		32

NATIONAL PERCENTILE

1 2 5 10 20 30 40 50 60 70 80 90 95 98 99

READING Vocabulary

READING Comprehension

LANGUAGE Mechanics

LANGUAGE Expression

LANGUAGE Spelling

ARITHMETIC Computation

ARITHMETIC Concepts

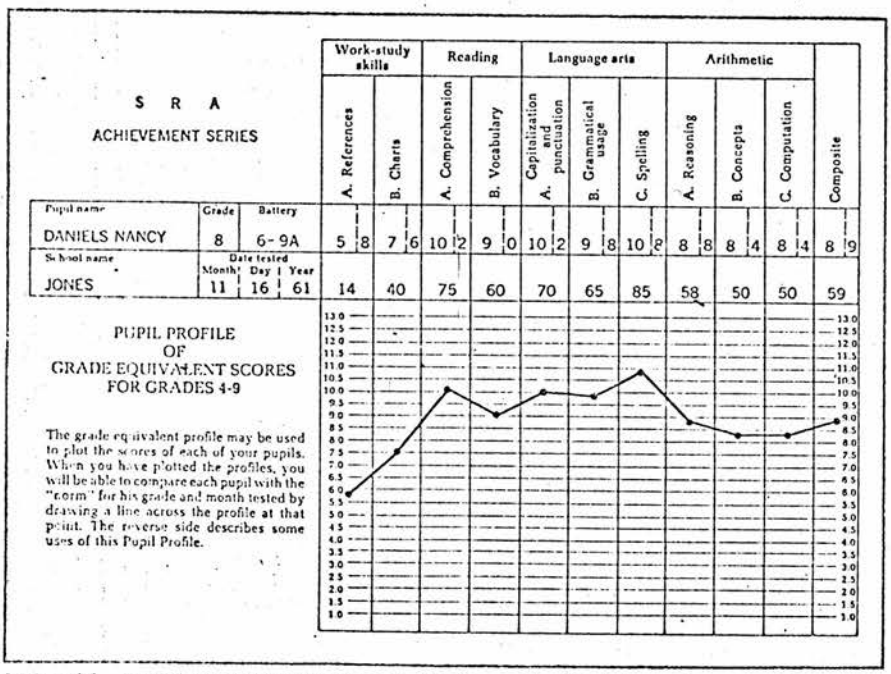
ARITHMETIC Applications

STUDY SKILLS Reference materials

STUDY SKILLS Graphic materials

Q2

FIGURE 42 An adaptation of the California Test Bureau pupil profile chart.
(California Test Bureau, *Comprehensive Tests of Basic Skills: CTB Individual Test*)



S R A
ACHIEVEMENT SERIES

Pupil name	Grade	Battery
DANIELS NANCY	8	6-9A
School name	Date tested Month	Day	Year
JONES	11	16	61

	Work-study skills		Reading		Language arts			Arithmetic			
	A. References	B. Charts	A. Comprehension	B. Vocabulary	A. Capitalization and punctuation	B. Grammatical usage	C. Spelling	A. Reasoning	B. Concepts	C. Computation	Composite
	5 8	7 6	10 2	9 0	10 2	9 8	10 8	8 8	8 4	8 4	8 9
	14	40	75	60	70	65	85	58	50	50	59

PUPIL PROFILE
OF
GRADE EQUIVALENT SCORES
FOR GRADES 4-9

The grade equivalent profile may be used to plot the scores of each of your pupils. When you have plotted the profiles, you will be able to compare each pupil with the "norm" for his grade and month tested by drawing a line across the profile at that point. The reverse side describes some uses of this Pupil Profile.

13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0

FIGURE 43. An adaptation of the Science Research Associates pupil profile chart.

items was divided into blocks of ten, and the mean score for each group obtained. Translated into a graph, this analysis permitted us to observe both the quality of performance and the rate of progress by reference to how the scores were distributed at different intervals during the reading. By averaging the scores for each of the groups in each block of ten items, we balanced out idiosyncratic responses while at the same time permitting each ability group to produce the norming pattern to be identified with its own performance.

Figures 44, 45 and 46 present the reading performance profiles for upper, lower and middle groups, only the upper being considered satisfactory in terms of proficiency level. Since the rate of item deletion was every tenth word it is fairly easy to identify say items (30 to 40) with roughly the 350 to 450 word mark as the first sentence of this text which was left unmutilated was about 50 words long. If our interest was in how subjects performed at about the 700 word mark, we could easily refer to performance in the (60 to 70) item block.

The three profiles presented here reveal some very interesting features that help to identify them.

READING PERFORMANCE PROFILE

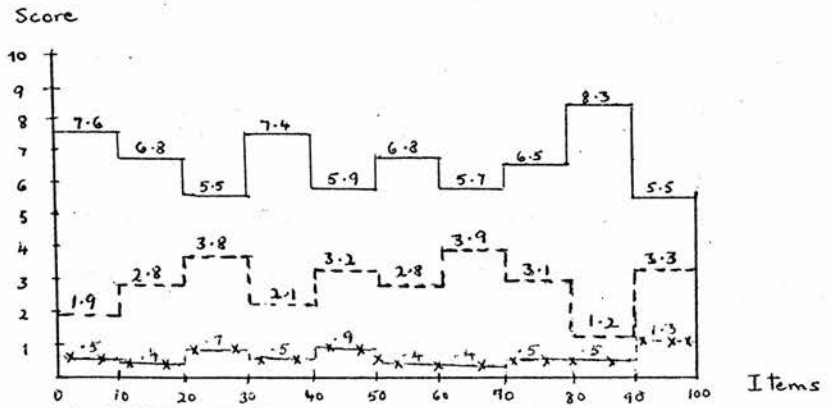


Fig. 44 UPPER THIRD PROFILE OF PROFICIENT READER BASED ON N= 63

—— Correct Score
- - - Incorrect Score
×—×—× Omission Score

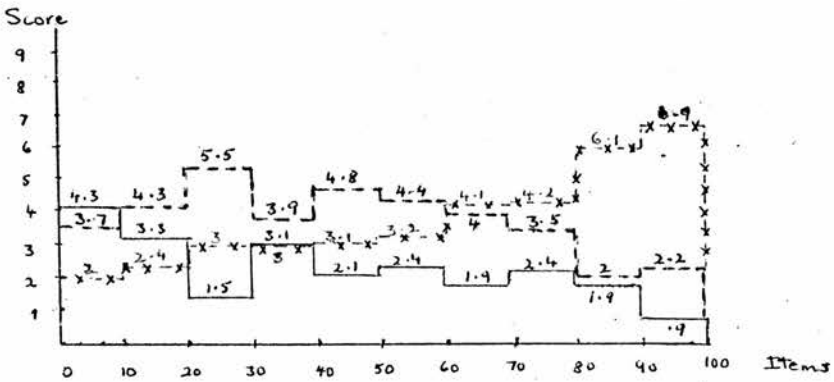


Fig. 45 LOWER THIRD PROFILE OF NON-PROFICIENT READER BASED ON N= 63

—— Correct Score
- - - Incorrect Score
×—×—× Omission Score

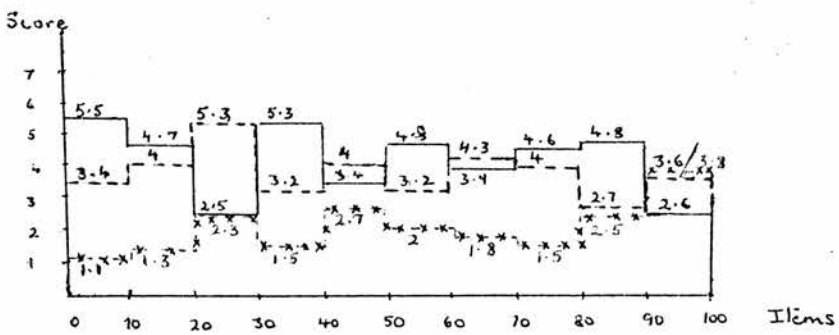


Fig. 46 MIDDLE THIRD PROFILE OF NON-PROFICIENT READER BASED ON N= 64

—— Correct Score
- - - Incorrect Score
×—×—× Omission Score

Fig. 44 presents the behavioural pattern of the proficient reader, which is clearly noted for its stability. With increased context there is a slight increase in the accuracy score (See 9.3.1, Table 59) towards the latter part of the cloze test, while the omission score remains negligible throughout, regardless of length.

The lower third profile in Fig. 45 shows how length affects performance. Whereas responses are distributed more between correct and incorrect at the beginning, with increasing context there is a decline in response output, and a marked increase in number of omitted items representing inability to retrieve any item that can conceivably fit the constraints. We see here the evidence of a breakdown in communication between writer and reader. As the profile indicates, the patterns for incorrect and omitted items diverge, with the gap between attempted responses and no responses, widening with length.

The middle third profile in Fig. 46 reveals a different trend. What occurred earlier in the test for the lower third group is evident towards the end of the test for this group, that is, the decline in correct responses and an increase in omitted responses. All lines converge in this profile, and it would suggest that the reader at this level is finding it difficult towards the end to maintain comprehension.

It would be interesting to see how the reader here performed in a test using texts longer than this span. The indications seem to suggest that his performance might

well reflect patterns indicative of a breakdown in communication if the text proceeded further with fewer attempted responses and an increasing number of omitted items.

We visualise the use of the reading performance profile in the following way:

(i) Administer Cloze Test based on Passage 1

(ii) Score correct responses using the acceptable word method (Details of acceptable words are given in Appendix 1.B.1.2, Page 14)

(iii) Score number of omitted responses

(iv) Obtain totals for each block of ten items

(v) Compare reader's profile with mean scores in Appendix 4.C.1, page 141, or from Figures 44,45,46 to see which type he represents.

It is likely that the reader will exhibit all the tendencies of a proficient reader for the first half, but may fail to complete the test clearly indicating lack of speed. Alternatively the breakdown, if gradual, would suggest that length of the text was causing increasing problems. Practice with longer texts would then be indicated. But we primarily see this reading performance profile as a diagnostic instrument providing more information about individual ability than could be gleaned from total scores. As we reported earlier the widening differences between proficiency levels were apparent around the (60 to 70) item block which represented the 650 to 750 word mark. Since of the ten items comprising this block, only two of the ten were

functional items,we investigated to see whether ease or difficulty was affected by the ratio of functional to content items in each block, the former having proved easier for all. [See page 163 in Appendix 4.D.1. for analysis.]

Figure 47 shows the number of functional items present in each block. Compared with this are the correct score results for the three groups. As is evident there were two blocks which had very few functional items. In the first one (20 to 30 item block), there was a marked decrease in correct scores for all readers suggesting a possible connection between number of content items and difficulty level in a cloze test. But this simple explanation was dispelled by a study of responses in the second block - the 60 to 70 item block. Here the connection was less evident, especially with the lower third group. In the (50 to 60) item block there are seven functional items but the average score for the lower third group was only 2·4. In the (60 to 70) item block when the number of functional items dropped to two out of the ten, the average score declined to 1·9, a marginal decrease. The influence then was negligible at this stage confirming the interpretation that with this group at least, declining performance is caused by many more factors since they are unable to take advantage of grammatic and syntactic cues to complete items that come from a closed system, especially when the text is long.

Ease and difficulty are not dependent on the number of functional items. But of interest is the parallel pattern

of correct scores of upper, middle and lower third groups up till the (60 to 70) item block. After this they diverge.

Textual difficulty as we have shown is not so easily identifiable.

We carried out a similar comparison between the distribution of functional items in the text and the distribution of omitted responses, once again to examine the possibility of any easily discoverable connection. Fig. 48 presents this, and once again although there is an increase of omitted responses in the (20 to 30) item block which has only two functional items, there appears to be little direct connection between the two. The omission rate continues to climb inexorably with length for the lower third group regardless of the number of functional or content items in each block. The rise is less steep for the middle third group.

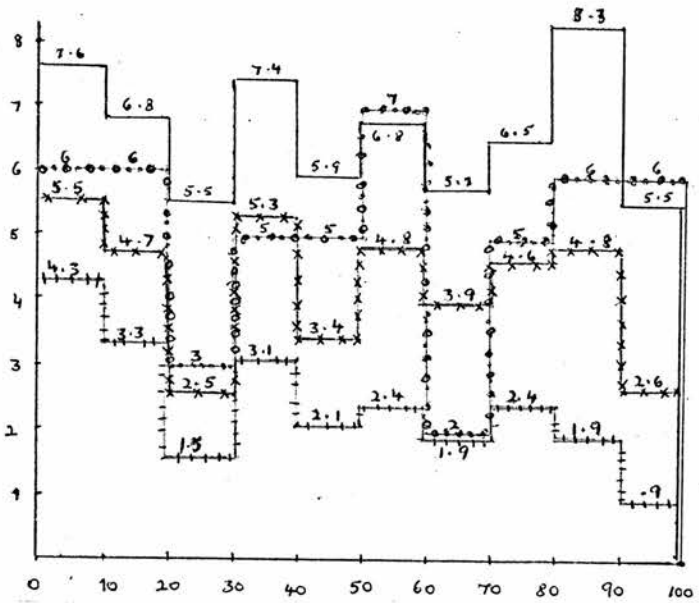


Fig. 47 DISTRIBUTION OF FUNCTIONAL ITEMS AND CORRECT SCORES

Number of Functional Items

Correct Score for Upper Third Group, N= 63
Correct Score for Middle Third Group, N= 64
Correct Score for Lower Third Group, N= 63

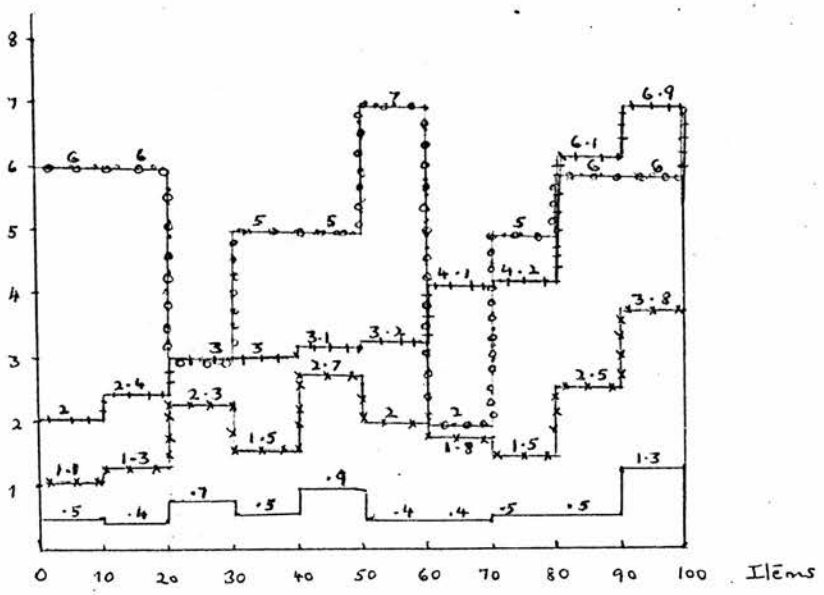


Fig. 48 DISTRIBUTION OF FUNCTIONAL ITEMS AND OMISSION SCORES

Number of Functional Items

Omission Score for Upper Third Group, N= 63
Omission Score for Middle Third Group, N= 64
Omission Score for Lower Third Group, N= 63

This then is the Reading Performance Profile, a descriptive and a diagnostic tool, of use to both tutor and non-native reader.

As we stated in the hypotheses investigated, cloze activates items in store and thereby reflects the level of linguistic proficiency. And if too many items are incorrectly supplied, or indeed fail to be supplied at all, the subject can be assumed to be deficit in the second language, and to this extent his reading ability, too, is inadequate for the demands imposed by academic instruction.

'When a person receiving the message fails to respond because of different conceptualisations of the verbal symbols, misinterpretation of language, or because of limitations in cognitive ability which prevent understanding, we conclude that there is no communication. In the case of communication in reading we say that the reader fails to comprehend' Kingston (1970:233). Our reader has been assumed to have adequate cognitive ability and the conceptualisation necessary for communication, though in his first language. His knowledge of the second language is however deficient if he fails to produce evidence of the understanding from printed input.

9.7 SUMMARY

In this chapter, results obtained from the task induced operations were discussed both independently in terms of the individual experiments, and with reference to the hypotheses postulated to validate the reading model. Six of the seven hypotheses were confirmed, one partially. The

seventh was not investigated because of faulty experimental conditions and although no data was produced that substantiated the hypothesis, we preferred to treat this as unconfirmed rather than incorrect.

The results of the cloze test were then used to develop the reading performance profile based on group mean scores representing rate of progress through the cloze passage for upper, middle and lower levels of ability.

In addition, performance scores and omission scores were presented in relation to the distribution of functional items in the text to examine whether ease or difficulty was affected by the number of items coming from a closed system since the latter had proved easier to complete. As was evident from the comparison, the effect of the number of functional items on successful completion was small, the results being influenced by a host of other variables subsumed under language deficit, and which affected the level of comprehension and reading ability in general.

CHAPTER TEN

THE MODEL REVIEWED

10. GENERAL REMARKS

In this chapter, the experimental data obtained will be used to review the reading model hypothesised earlier and suggest further ways in which it could be useful.

Also included will be some suggestions that based on the model and our findings could be used to improve the quality both of instruction and evaluation of the ability being reviewed.

10.1 THE READING MODEL REVIEWED

Reading as we viewed it was a multi-level process ranging from overt observable behaviour to unobservable cognitive processing. Since we were primarily interested in eliciting reliable evidence of reading ability from which we could infer the reader's competence, one of the goals of our reading model was the formulation of theoretical accounts of internal processing and the specification of various components, mechanisms and activities that were together capable of accounting for reading behaviour. Emphasis was placed on how understanding might be achieved, what it indicated and why it varied from one individual to another. However since the model as a model would have remained purely theoretical and speculative, we developed

hypotheses that by use of the systems, components, capacities and processes postulated, could be experimentally investigated.

The need for eliciting measurable response led to a careful analysis of input in terms of difficulty level and the mode and manner of presentation on the one hand, and a detailed specification of how the responses generated were to be evaluated and interpreted on the other. The attention therefore shifted from the cognitive processing to observable behaviour, and the type of behaviour elicited was controlled by the demands imposed by the tasks set that enabled the specific hypotheses to be investigated.

The problem as we saw it was that we expected from responses to task-induced operations, the elicitation of data that would enable us to infer the presence and use of cognitive processes in a readily quantifiable manner. Reading itself does not possess such identifiable characteristics, hence the need to set up conditions that would permit inferential judgements which, though based on observable output, were both reliable and valid.

By attempting to control the various task variables that we had identified, we focussed our attention on the dependent variable which we recognised was not the single variable so often isolated in specific experimental studies. In fact as we saw it, the inflexible law of the single variable could not be applied here if, as we postulated, reading ability was dependent not just on what factors were associated with the task, but with factors outside the control

of the tester and not circumscribed by the testing situation. The two major independent variables the reader broughtto the task were the level of linguistic proficiency in the second language, and the skills and strategies he already possessed, both of which interacted with input to produce the output we would evaluate.

The only way some kind of control could be exercised was by the choice of techniques peculiarly suited to the activation of one of the variables more than the other, although continual interaction was present. By using measures for activating and simulating a search process by a cloze-type task, we obtained evidence of linguistic proficiency supported by individual strategies. As our results showed, the high correlations obtained between the cloze test, and the standardised test ELBA especially, confirm this view.

On the other hand the choice of techniques that necessitated the exercise of various skills and strategies to solve or find solutions to problems permitted us to focus on the second variable. This was one of the major advantages of postulating a model that permitted identification of what were clearly two aspects or facets of a very complex ability. The obvious inference was that if non-native subjects were to read efficiently, possession of knowledge of linguistic items alone was insufficient. Not only did they need to know the use, i.e. the function of language forms, but they also needed to exercise various skills and strategies in abstracting the meaning from the text. In this connection let us briefly examine how our readers

could be categorised if they lacked the necessary skills.

Oaken, Wiener and Cromer (1971) had suggested that reading disability could be explained in terms of 'deficit' or 'difference'. Our readers, at least the non-proficient readers, had betrayed a distinct deficit in terms of linguistic items and forms. This is however not unexpected with second language learners. If the deficit was very marked, as it was with the lower third group used in the profile, it was obvious that they would be unable to benefit from the course of studies they were enrolled on without intensive remedial tuition. If however the deficit exhibited was similar to that of the middle third group in our reading performance profile, the development of effective strategies would help to compensate somewhat for this lack of linguistic ability. In fact as we suggested during the discussion of our model, the use of individual skills and strategies are an essential attribute of the proficient reader if he is also efficient. By proficient we have throughout the thesis meant linguistic adequacy, whereas by efficiency we have meant maximum comprehension with maximum economy of time and effort.

One of the problems with printed input is that when difficult concepts are presented or when informational density is high, this is often reflected in the use of complex writing patterns that help to give cohesion and coherence to the whole thought. The difficulty in processing then, reflects the reader's ability to recognise the relevant syntactic groupings or chunkings that make for

ease of comprehension.

In the spoken mode, the speaker uses intonational features, reduced speed and pauses, to chunk his words and ideas for greater clarity. In the written mode, however, it is the reader who must ensure that the largest manageable unit - the LMU - coincides with syntactic boundaries so that meaningful chunks can be processed enabling him to read with understanding. With reduced speed the LMU selected by the poor reader may be so minimal that higher level processes cannot operate, and meaning beyond the immediate constituent level becomes more difficult.

'It appears that there is simultaneous consideration of one or more words which perform syntactic or lexical function or both, but that there is a hierarchical process by which lower level units are synthesised into higher level units. Thus all readers use simultaneous and successive synthesis at some points', Jenkinson (1973:160).

With our NNS we recognise that at specific intervals or choice points the linguistic options may not be available in the foreign language. If the language user is to actively participate in the process of communication he must evolve a processing system which will permit him to conceptualise his experience when necessary in English. He must be able to draw inferences from relevant cues which must be recognised and identified as being relevant. Since in cloze, the NNS is required to produce evidence that the reading is being intelligibly undertaken, the inability to produce the odd item is unimportant and may only be proof of non-possession

of the necessary item in the second language. If however the omitted items increase in quantity there is every reason to assume that even if comprehension may have occurred, there is objective evidence that associated vocabulary and meaning forms are absent in English reflecting a stage of insufficient knowledge of the language needed. Linguistic insufficiency or language deficit can become a major obstacle to success in academic work, so that there was every justification for interpreting the omission scores as failure, if not incomprehension. The foreign learner in reading must rely on the linguistic rather than the perceptual context, a task made more difficult since the context to be recreated is itself an alien factor to be also learned.

The deficit model then did provide an explanation for many of the NNS who took part in the experiments. But a question that was raised though not answered needs consideration here. To what extent did a 'difference' model contribute to linguistic inadequacy for the NNS? If the latter had been exposed to only one kind of English, the English of second language textbooks, did this limited exposure render it difficult for him to comprehend English in natural contexts? If there was a problem, there were two ways in which this problem could be solved. Either the reader had to learn to adapt to new forms, or the printed input had to be adapted to suit the reader. In our review of the literature we had briefly described one such experiment where input was re-written to match the reader's language [See 3.4.3, Peltz (1974)].

This choice or even possibility is not open to our NNS. He must quickly adapt to new demands made by forms and styles he is unfamiliar with. Help here can come through inculcation of strategies. In our thesis, strategies have been viewed as techniques or approaches to tasks which the reader either evolves through use or via explicit instruction. They can be taught, practised and learnt. They are not 'inherent' in the task as Cooper and Petrosky (1976:194) suggested. According to them strategies were 'natural abilities natural to the reading process'.

It was because of our interest in these skills and strategies so essential for efficient reading, that we set such store by the taxonomy of abilities that we attempted to activate and so measure in the multiple-choice test. The classification of items according to type of ability being measured was obviously an oversimplification of the processes utilised for the solution to the questions.

There is always the danger that observed performance becomes equated with the ability we infer is being measured. But in all three versions of the multiple-choice test there was strong evidence to support the conclusion that questions involving inference and evaluation were difficult for all readers in general; and questions on detail were easier than questions involving more generalised concepts.

This can only be meaningfully interpreted in terms of the cognitive processes involved which we referred to in our model. When factual information is explicit in the text the demands on the reader entail location of this information,

and matching it with the distractors. When factual information is implied, the reader must draw a logical inference and then match. However with the information not readily available but having to be inferred by reading between and beyond the lines, the demands on memory and cognitive activity increase.

As we saw in Chapter Five, attention and processing capacity are limited. When this capacity must be divided between selection and storage of information on the one hand, and cognitive activity with the information on the other, before matching with the item distractors can occur, item difficulty increases. The reader who has little or no difficulty with selection of relevant information from the text can pay more attention to reasoning than the reader who overloads memory capacity with verbal data because he cannot effectively use critical skills. Lower level analysis takes up his attention and higher level processes cannot operate.

As our model made clear, reading successfully necessitated parallel processing. While perceptual processes were making maximal use of central and peripheral visual cues aided by syntactic recognition of phrase boundaries, semantic processing was simultaneously occurring. With the continual feedback from working memory, the level of predictability ought in theory to have increased as 'redundancy utilisation' (Weaver and Kingston, 1963:259) assumed greater importance. Although there was some evidence obtained in a slight increase in successful scores

for the proficient reader towards the latter part of the long text, this accumulation of information needs to be investigated more carefully. It was partially investigated by the summary-writing experiment, but the focus there was primarily on the relation between process and product.

A very brief analysis of some of the kinds of responses generated for the cloze task also helped in confirming our explanations of how cognitive processes functioned when NNS were presented with verbal input with which they were not very familiar. This analysis is provided in Section 10.1.1.

It was evident that a knowledge of the grammar and syntax of English would be of particular help in the completion of deleted items that belonged to a closed system. That is, functional items would be easier since the choice here was confined to a very limited paradigm and therefore retrieval mechanisms would process that much faster. This was partly confirmed by the higher overall score for functional as opposed to content items. In some cases the choice would be reduced to one so that with the proficient reader, predictability was raised to 100%, and retrievability automatic.

With content items, because they belonged to an open class, choice was constrained by appreciation of the contextual restraints. Of particular interest were the various strategies employed by readers to complete difficult items. Some readers countered by retrieving words from neighbouring contexts, others used these as stimulus cues

to recall words associated with them in other contexts, although they were inappropriate in the context of the test.

There were also instances where rules regarding word formation were applied to produce English-type items, and in some instances words from the first language were used.

Of particular interest were the very many instances of unilateral context influencing the selection so that the item was incorrect in the bilateral situation. This reflected the decreasing influence of context when items could not be easily retrieved and attention narrowed to the immediate neighbourhood.

10.1.1 Brief Analysis of Incorrect Responses

In the following examples, the correct word is presented in the underlined space, and the various alternatives suggested come after it.

(a) Inappropriate words used from preceding context and LTM:

(i) cattle, sheep, goats and pigs were used. . .
Words suggested : cows, dogs, fish

(ii) the use of the horse, ass, ox and camel for transport.
Words suggested : horse, cows, sheep

(iii) When the basis of city life was established, the urban tradition was drawn from the ancient cities of the Middle East.
Words suggested : basis, town

(iv) the potter, the weaver and the flint polisher
Words suggested : shoe, car, furniture, wood, rice

(b) Failure to recognise idiommatic usage:

(i) the fly in the ointment

Words suggested : of, on, from

(ii) for the sake of convenience

Words suggested from lexical field of the passage : law, way, jurisdiction, result, rule

(iii) no one knows for sure

Words suggested : very, rather, it

(c) Instances of unilateral context:

(i) a community of millions is beyond human . . .

Words suggested : people, inhabitants

(ii) many of these

Word suggested : times

(iii) a greater and more reliable supply . . .

Words suggested : result, deal, food, the, amount

(iv) but in doing so it made possible a . . .

Words suggested : much, had, was

(d) Words with some similarity of form and meaning:

For pedestrian	: Words suggested:	walk, walking
Renaissance	:	renaiscientist, renocentist
Ages	:	adge
removed	:	disappeared, tented
it (meaning history)	:	historician
exhaustion	:	exhoistment, exhaust
led	:	leaded

This is an area which might well prove worth more detailed investigation. Of some 60 papers selected at random, only 20 had instances where more than two (of the 100) items

were altered. Of the total of 80 such instances, 48 resulted in a correct score; in all cases where two or more alternatives were given all were recorded, but the last one only, evaluated. Very rarely were there instances of correct responses being changed to become unacceptable. This does suggest that the time provided was sufficient for an initial reading but not enough for re-evaluation of the test after completion. The results do seem to provide evidence of understanding during the reading process at first exposure.

10.1.2 Speed of Information Integration

In this study, we were concerned not with speed as a variable to be measured, since this was of itself meaningless for our purposes, but as a factor contributing to reading efficiency. Our analysis gave us an accuracy score and efficiency was judged by the number of items accurately completed in the cloze test, or answered in the multiple-choice test or via the number of ideas recalled in the summary task within the time limit set. This is also why we attach so much value to the reading performance profile since it enables us to judge whether an individual lacks overall linguistic competence or, if he is a proficient but slow reader.

In many experiments using good and poor readers the grouping was, as our review in Section 3.6. indicated, made on the basis of other standardised tests. In our thesis the division into different proficiency levels was made on the basis of the tests prepared, and the reliability and

validity of the decisions confirmed by statistical measures including correlation with standardized batteries which served as external criteria.

The pragmatic validity of six of the hypotheses investigated enabled us to affirm with a fair degree of confidence, that the reading model postulated was supported by experimental evidence.

10.2 SUGGESTIONS FOR FURTHER RESEARCH

Our concern has been to increase the generalizability and predictability of the model across content, population and other variables. Although we used five long texts and four shorter ones, population samples in Britain and overseas, and three techniques, there is obviously need for further experimentation to confirm the validity of inferences made on the basis of our results.

We present below specific areas in which we would like to see further research undertaken.

(i) Not only should other materials be used to support our findings but, more importantly, the investigation needs to concentrate on variations in length. Since this has far-reaching implications especially for the population we were concerned with, we would suggest that texts of both 1000 words and about 2500 words be used using the cloze technique to see whether the effects of an "attention span", if it exists, occurs in a recurring cycle at fairly stable intervals. Self-contained texts would be particularly appropriate

so that information content from one part could affect comprehension across spans.

Specific questions to be asked would include the following:

(a) Is "attention span" a stable, and therefore easily predictable, phenomenon?

(b) Or does it vary with difficulty level of text and reading efficiency level of the subject population?

(c) Is it associated with amount of attention and processing effort i.e. degree of concentration expended by the reader and therefore dependent on factors like text, reading ability, reading objectives etc., or a time factor?

(ii) As reported earlier we failed to examine the influence of location of distractors on item difficulty. We see this as an important area since it is to some extent associated with text structure. Information related to one item or incident often occurs in non-sequential form, and ability to abstract content from different parts of the text for evaluation and inference is necessary for understanding.

(iii) We would like to further investigate the validity of our model by using native speakers who already possess the linguistic proficiency necessary for successful reading. We suggest they could be required to process long texts at academic level. Variations in performance would then be the result of strategies and skills which

could be further investigated by the setting up of specific tasks that made increasing demands on the cognitive processing system. We view the organisation of the accumulation of information as probably one of the best evidences of reading efficiency.

Additional insights into strategies used by the reader to process text could be achieved by an analysis of errors obtained *via* cloze. These could be classified according to probable cause of error such as we briefly attempted in the last section. The use of analytic procedures such as were utilised in the thesis would permit comparison of the reader's performance at different points in the text. We need more information regarding the interaction between reader and text, and understanding of textual organisation might well prove to be an important variable affecting the difficulty level of texts. As we saw, readability measures, deletion and lexical counts failed to account for variations in performance of the subjects in our study when faced with different passages.

(iv) One of the more interesting offshoots of the experimental study was the 'reconstruction-of-text' experiment using an abbreviated version as stimulus. By providing varying amounts of text - either through cloze, or through notes or questions of the multiple-choice type - for recomposition of the message, both reading ability and organisational strategy would be reflected in the success with which subjects imposed a coherent frame of reference and abstracted information from the limited

data. This would also provide evidence of productive ability in the second language, an area which has tended to be neglected for reasons associated with administrative and practical convenience.

This investigation of both the reading process and the reading product along with the advantages of the various testing and analytic techniques has permitted us to obtain interesting evidence regarding possible explanations of processes involved in the comprehension activity. By integrating suggestions and experimental findings into a comprehensive framework we have attempted to focus on explicit areas of further research.

10.3 APPLICATIONS OF FINDINGS

Of particular relevance is the finding that length is a factor that contributes to differentiation between levels of reading efficiency. Nor can a proficiency level based on the ability to read short passages be automatically assumed to apply to longer texts such as are used on academic courses.

Also of direct consequence is the confirmation that inferential and evaluative skills make greater demands than reading for factual detail explicitly stated in the text. The implications are obvious. Because length can create a problem especially when the critical faculty must be exercised, practice is needed. Skills cannot be assumed to exist especially with the non-native speaker since modes of thinking and therefore expression are not universal across language and culture.

Our model has been an information processing model in which input has been broken down or built up as necessary for the identification of syntactic and semantic units. As we have tried to show, because of the interactive nature of the process 'the total act of reading involves almost the whole human brain, for we get out of reading very largely what we bring to it', Mackworth (1972:703).

There are a number of important skills and concepts essential to reading which vary from one culture to another and which are not included in courses of instruction either because the teacher is unaware of their existence and importance, or because it is assumed that subjects already have them.

A course of instruction that attempted to teach strategies would make for efficient reading. Familiarity with conventions of lay-out, the use of content and section headings, indentation, size of print etc., are aimed at encouraging the reader to perceive the text in more than one dimension since all these clues can help in processing strategy.

So too is recognition that different kinds of texts often use different kinds of formats. Journals require information to be concise, conforming to a set order, where an abstract of the article is followed by the general problem, related literature, specific hypothesis, experimental procedure, results and discussion. Books especially text-books are often sequentially ordered unlike collections where articles are classified according to content and theme.

These are aids to the reader enabling him to focus more attention on one part of the text than another. We found in the course of using lexical counts to determine difficulty levels, that earlier parts of texts tended to have a higher lexical variation count. More new words and therefore possibly ideas were introduced at the beginning than later, so that the reader needed to slow down, spending more time at the beginning and speeding up towards the latter half rather than spending the same amount of time and energy throughout the text. Our investigation has therefore suggested both areas for further research, and the application of our findings to courses of instruction in English as a second language.

10.4 SUMMARY

We recognise that we have not followed the current experimental style where specific small experiments are designed to examine specific questions. We chose instead an alternative approach that as Newell said (1973:303) is to 'focus a series of experimental and theoretical studies round a single complex task, the aim being to demarcate that one has a sufficient theory of a genuine slab of human behaviour'. This appears to be the only way to integrate information and knowledge rather than add to the diversity of new interesting studies that remain isolated.

This, in essence, is what we have tried to do.

CHAPTER ELEVEN

CONCLUSION

This thesis has been primarily concerned with measures and techniques aimed at evaluating the reading efficiency of non-native speakers who for various reasons undertake courses of instruction at academic level in a second language.

Since so little is known of the complex cognitive activity involved in reading, we adopted the view that an attempt to evaluate the degree of skill with which the reader interacts with the text at first exposure, will provide us with a measure of ability related to the reading process. A technique that seemed particularly suited to this task was 'cloze'. Information regarding the process stage would be taken as providing complementary evidence to that obtained by the utilisation of techniques more directed towards the reading product.

To provide the very necessary rationale for the stand taken, we presented a five-stage reading model for use with English as a second language. Based on previously reviewed literature, the reading model functioned 'as a method of specifying the components with which we must deal and [enabled us to generate] testable hypotheses,' Clymer (1968:12). It forced us to organise facts and data obtained from experimental findings into a rational framework, and helped us to give construct validity to

the techniques we selected to activate the processes and skills we aimed to measure.

We used our information-abstracting model to explain how variations in reading ability were a function of two major factors interacting with input. These were the linguistic proficiency of the reader, and the skills and strategies he needed to exercise in order to read efficiently for academic work.

The results of this research may be summarised as follows:

1. Process and product are complementary approaches to an investigation of reading ability, both equally necessary for an understanding of this complex activity.
2. Process can be regarded as the reader's interaction with the text, and a technique that seems peculiarly suited to evaluate this is 'cloze'.
3. A 'cloze' test given under timed conditions that measure reading comprehension at first exposure, provides us with a reliable and valid indices of linguistic proficiency if it satisfies two conditions. One is that the textual material used is of the appropriate level of difficulty for the population and, second that it contains a large number of deleted items. The Cloze test based on Passage 1, and tried out with a fairly large sample, has been demonstrated as being a reliable and valid instrument for this purpose, and has been used to infer the degree of

reading competence the subject brings to the testing situation.

4. The Reading Performance Profile shows how performance at this test can be analysed to provide more information regarding the subject's reading ability. The analytic approaches to the data show the value of a technique that permits us to obtain evidence regarding rate of progress and distribution of scores at intervals in the test, so that its diagnostic value can be utilised for remedial purposes.

5. The absence of correlation between the summary task and the proficiency test, ELBA, suggests that there may be a need for requiring some kind of evidence reflecting the productive competence of the NNS in English. But this would have to be verified by more reliable testing.

6. The use of a taxonomy has confirmed the value of concentrating at test level on questions that necessitate the exercise of inferential and evaluative skills.

7. And perhaps, most importantly of all, the finding that length is a factor to be taken into account when assessing the reading efficiency of NNS readers at this advanced level.

In conclusion, we would like to suggest that the advantage of this thesis lies in the integrated nature of

the research, with experimental work supported by the reasoned explanations presented for consideration through the formulation of the working model, the latter representing in simplified fashion the reading process.

BIBLIOGRAPHY

1. Allen, J.P.B. and Alan Davies, ed. (1970). Testing and experimental methods, ECAL vol. 4. London : Oxford University Press.

2. Anderson, John (1970). A technique for measuring reading comprehension and readability, English language teaching, XXV, 178-182.

3. Anderson, R. and David Ausubel, ed. (1966). Readings in the psychology of cognition. New York : Holt, Rinehart and Winston Inc.

4. Bach, E. and R.T. Harmes, ed. (1968). Universals in linguistic theory. New York : Holt, Rinehart and Winston Inc.

5. Baratz, Joan (1969). A bi-dialectal task for determining language proficiency in economically disadvantaged Negro children, Child Development, 40, 889-901.

6. Bartlett, Sir Frederick (1932, 1950). Remembering : A study in experimental and social psychology, Cambridge : Cambridge University Press.

7. Berger, Natalie S. (1975). An investigation of literal comprehension and organizational processes in good and poor readers. Unpublished Ph.D. thesis. University of Pittsburgh.

8. Berlo, David (1960). The process of communication : An introduction to theory and practice, Washington : Holt, Rinehart and Winston Inc.

9. Bleismer, E.P. and A.J. Kingston, ed. (1961). Phases of college and other adult reading programs. Milwaukee : N.R.C. 10th year-book.

10. Bloom, Benjamin S. ed. (1956). Taxonomy of educational objectives : The classification of educational goals. Handbook 1 Cognitive domain. New York : McKay.

11. Bloom, B.S. Hastings, J.T. and G. Madaus (1971). Handbook for formative and summative evaluation of student learning. New York : McGraw Hill Book Co.

12. Bloomfield, L. (1933). Language. New York : Holt, Rinehart and Winston Inc.

13. Bond, Guy and Miles Tinker (1957). Reading difficulties : their diagnosis and correction. New York : Appleton-Century-Crofts Inc.

14. Bormuth, John (1968). The cloze readability procedure, Elementary English, 45, 429-436.

15. Bormuth, John (1969a) Factor validity of cloze tests as a measure of reading comprehension ability, Reading Research Quarterly, IV, 358-365.

16. Bormuth, John (1969b) An operational definition of comprehension instruction. In Goodman and Fleming (1969 : 48-60).

17. Bormuth, John. Manning, Carr and Pearson (1970). Children's comprehension of between and within sentence syntactic structures, Journal of Educational Psychology, 61, 349-357.

18. Burkart, Kathryn N. (1945). An analysis of reading abilities, Journal of Educational Research, 38, 430-439.

19. Carroll, J. Carton, A. and C. Wilds (1959). An investigation of 'cloze' items in the measurement of achievement in foreign languages. Cambridge, Massachusetts : Research and Development Reports : College Entrance Examination Board.

20. Carroll, John B. (1971). Learning from verbal discourse in educational media : a review of the literature. Princeton : Educational Testing Service Research Bulletin 71-61.

21. Carroll, J. and R. Freedle, ed. (1972). Language comprehension and the acquisition of knowledge. Washington : V.H. Winston and Sons.

22. Carver, Ronald (1974). Improving reading comprehension : Measuring readability. Final report. Maryland : American Institutes for Research.

23. Chafe, Wallace (1970). Meaning and the structure of language. Chicago : Chicago University Press.

24. Chafe, Wallace (1972). Discourse structure and human knowledge. In Carroll and Freedle (1972 : 41-69).

25. Chall, Jeanne (1958), Readability : An appraisal of research and application. Columbia : Ohio State University.

26. Chase, William, ed. (1973). Visual information processing. New York : Academic Press.

27. Chomsky, Noam (1965). Aspects of the theory of syntax. Cambridge : M.I.T. Press.

28. Cicourel, A.V. (1973). Cognitive Sociology. Harmondsworth.

29. Clymer, T. (1968). What is reading : some current conceptions of reading. In Robinson (1968 : 7-29).

30. Conrad, R. (1972). Speech and reading. In Kavanagh and Mattingly (1972 : 205-240).

31. Cooper R. and A. Petrosky (1976). A psycholinguistic view of the fluent reading process, Journal of Reading, (December) 184-207.

32. Corder, S.P. and E. Roulet, ed. (1973). Theoretical linguistic models in applied linguistics. Brussels : AIMAV.

33. Darnell, Donald K. (1968). Clozentropy : A procedure for testing English language proficiency of foreign students, Speech Monographs, 36-46.

34. Davies, Alan, ed. (1975). Problems of language and learning, London : Heinemann.

35. Davies, Alan (1977). The construction of language tests. In Allen and Davies (1977 : 38-104).

36. Davis, F.B. (1944). Fundamental factors of comprehension in reading, Psychometrika, 9, 185-197.

37. Davis, F.B. (1968). Research in comprehension in reading, Reading Research Quarterly, III, 499-545.

38. Davis, F.B. ed. (1971). The literature of research in reading with emphasis on models. New Jersey : Graduate School of Education, Rutgers University.

39. Davis, F.B. (1972). Psychometric research on comprehension in reading, Reading Research Quarterly, VII, 628-678.

40. Di Vesta, F.J. and G.S. Gray (1973). Listening and note-taking : II Immediate and delayed recall as functions of variation in thematic continuity, note-taking and length of listening review intervals, Journal of Educational Psychology, 64, 278 - 287.

41. Downie, N. and R. Heath (1965). Basic statistical methods, New York : Harper International Student Reprint.

42. Farr, Roger ed. (1970). Measurement and evaluation of reading. New York : Harcourt, Brace and World Inc.

43. Fillenbaum, S., Jones L.V. and A. Rapoport (1963). The predictability of words and their grammatical classes as a function of rate of deletion from a speech transcript, Journal of Verbal Learning and Verbal Behaviour, 2, 186-194.

44. Fillmore, C. (1968). The case for case. In Bach and Harmes (1968).

45. Frase, L.T. (1972). Maintenance and control in the acquisition of knowledge from written materials. In Carroll and Freedle (1972 : 337-357).

46. Frederiksen, Carl H. (1972). Effects of task-induced cognitive operations on comprehension and memory processes. In Carroll and Freedle (1972 : 211-245).

47. Garfinkel, H. (1967). Studies in ethnomethodology. Englewood Cliffs, New Jersey : Prentice Hall Inc.

48. Geyer, John (1966, 1970). Models of perceptual processes in reading. In Singer and Ruddell (1970 : 47-94).

49. Gilliland, John (1972, 1975). Readability, London : University of London Press.

50. Gleason, H.A. (1965). Linguistics and English grammar. New York : Holt, Rinehart and Winston.

51. Gomulicki, B.R. (1956). Recall as an abstractive process, Acta Psychologica, 12, 77-94.

52. Goodman, K. and J. Fleming, ed. (1969). Psycholinguistics and the teaching of reading. Newark, Delaware : International Reading Association.

53. Goodman, K.S. (1969, 1970). Reading : a psycholinguistic guessing game. In Singer and Ruddell (1970 : 259-272).

54. Goodman, K.S. (1969). Analysis of oral reading miscues : Applied psycholinguistics, Reading Research Quarterly, V, 9-30.

55. Gough, Philip (1972). One second of reading. In Kavanagh and Mattingly (1972 : 331-358).

56. Greene F.P., ed. (1972). Investigations relating to mature reading. Milwaukee : National Reading Conference.

57. Guthrie, John T. (1973). Models of reading and reading disability, Journal of Educational Psychology, 65, 9-18.

58. Ingram, Elizabeth (1977). Basic concepts in testing. In Allen and Davies (1977 : 11-37).

59. Halliday, M.A.K. (1975). Talking one's way in : A sociolinguistic perspective on language and learning. In Davies (1975 : 8-26).

60. Halliday, M.A.K. and R. Hasan (1976). Cohesion in English. London : Longman.

61. Hasse, Karl D. and Richard Smith (1975). The effect of content and process questions written at memory and evaluation levels on 9th grade students' comprehension of a passage when the questions have been inserted either before or after segments of the passage. ERIC 095 510.

62. Hochberg, Julian (1970). Components of literacy : Speculations

and exploratory research. In Levin and Williams (1970 : 74-89).

63. Hochberg, Julian and Virginia Brooks (1970). Reading as an intentional behaviour. In Singer and Ruddell (1970 : 304-314).

64. Holmes, D.L. and Frank Smith (1973). Letter, word and meaning identification. In Smith (1973 : 50-69).

65. Hunt, Kellogg W. (1966). Recent measures in syntactic development. In Lester (1973 : 179-192). Printed earlier in Elementary English (1966) 43, 732-739.

66. Jakobson, Roman (1973). Main trends in the science of language. London : George Allen and Unwin.

67. Jenkinson, Marion D. (1973). Meanings in reading. In Karlin (1973 : 156-162).

68. Kaplan, Robert (1972). The anatomy of rhetoric : Prolegomena to a functional theory of rhetoric. Philadelphia : Centre for Curriculum Development.

69. Karlin, Robert, ed. (1973). Reading for all : Proceedings of the Fourth I.R.A. World Congress. Newark, Delaware : International Reading Association.

70. Kavanagh, J. and I. Mattingly, ed. (1972). Language by ear and by eye. Cambridge : M.I.T. Press.

71. King, David J. (1960). On the accuracy of written recall : A scaling and factor analytic study, Psychological Record, 10, 113-122.

72. Kingston, Albert J. (1960). A conceptual model of reading comprehension. In Bleismer and Kingston (1961).

73. Kingston, Albert J. (1970). The measurement of reading comprehension. In Farr (1970 : 230-236).

74. Klare, George (1974). Assessing readability, Reading Research Quarterly, X, 62-102.

75. Kolers, Paul (1970). Three stages of reading. In Levin and Williams (1970 : 90-118).

76. Krech, David (1973). Psychoneurobiochemeducation. In Lester (1973 : 1-16). Printed earlier in California Monthly (1969) LXXIX, 14-20.

77. Lester, Mark, ed. (1970, 1973). Readings in applied transformational grammar. New York : Holt, Rinehart and Winston.

78. Levin, H., Gibson, E. and J. Gibson, ed. (1968). The analysis of reading skill : A program of basic and applied

research. Final report. U.S. Dept. of Health, Education and Welfare.

79. Levin, H. and D. Jones (1968). Studies of oral reading XIII. Filled inter-word spaces and eye voice span (EVS). In Levin, Gibson and Gibson (1968 : 284-302).

80. Levin, H. and J. Williams, ed. (1970). Basic studies in reading. New York : Basic Books.

81. Lewis, D.G. (1967, 1972). Statistical methods in education. London : University of London Press.

82. Linnarud, Moira (1975). Lexis in free production : An analysis of the lexical texture of Swedish students' written work. Swedish-English Contrastive Studies Report No. 6. Dept. of English, Lund University.

83. Lunzer, Eric (1975). The effective use of reading. Paper presented to the UKRA Conference, Manchester 1975. A Schools Council Project.

84. Macginitie, W.H. (1961). Contextual constraint in English prose paragraphs, Journal of Psychology, 51, 121-130.

85. Mackworth, J. (1972). Some models of the reading process : learners and skilled readers. Reading Research Quarterly, VII, 701-733.

86. Mehler, J. (1963). Some effects of grammatical transformation on the recall of English sentences, Journal of Verbal Learning and Verbal Behaviour, 2, 346-351.

87. Melnik, A. and J. Merritt, ed. (1972). Reading today and tomorrow. London : University of London and Open University Press.

88. Miller, G.A. (1962). Some psychological studies of grammar, American Psychologist, 17, 748-762.

89. Miller, G.A. (1966). The magical number seven plus or minus two. In Anderson and Ausubel (1966 : 241-267). Printed earlier in Psychological Review (1956) 63, 81-97.

90. Miller, G. (1973a). The psycholinguists. In Lester (1973 : 17-35). Printed earlier in Encounter (1964) 23, 29-37.

91. Miller, G. (1973b). Some preliminaries to psycholinguistics. In Smith (1973 : 10-20).

92. Miller, G. and P. Johnson-Laird (1976). Language and perception. Cambridge : Cambridge University Press.

93. Miller, G. and J. Selfridge (1966). Verbal context and the recall of meaningful material. In Anderson and Ausubel (1966 : 268-280). Printed earlier in American Journal of

Psychology (1950) 63, 176-185.

94. Murray, D.J. (1967). The role of speech responses in short-term memory, Canadian Journal of Psychology, 21, 263-276.

95. Newell, Alan (1973). You can't play 20 questions with nature and win : projective comments on the papers of this symposium. In Chase (1973 : 283-308).

96. Norman, Donald A. (1968). Toward a theory of memory and attention, Psychological Review, 75, 522-536.

97. Oaken, R., Wiener, M. and W. Cromer (1971). Identification, organization and reading comprehension, Journal of Educational Psychology, 62, 71-78.

98. Oller, John W. (1971). Expectancy for successive elements. Address at TESOL Conference.

99. Oller, John W. (1973). Cloze tests of second language proficiency and what they measure, Language Learning, 23, 105-118.

100. Oller, John W. and Christine Conrad (1971). The cloze technique and ESL proficiency, Language Learning, 21, 183-195.

101. Olson, David R. (1972). Language use for communicating, instructing, and thinking. In Carroll and Freedle (1972 : 139-167).

102. Otto, Wayne (1972). Thorndike's 'Reading as reasoning' : influence and impact. In Melnik and Merritt (1972 : 37-43). Printed earlier in Reading Research Quarterly, (1971), VII, 435-442.

103. Peltz, Fillmore K. (1974). The effect upon comprehension of repatterning based on students' writing patterns, Reading Research Quarterly, IX, 603-621.

104. Perfetti, Charles (1969). Lexical density and phrase structure depth as variables in sentence retention, Journal of Verbal Learning and Verbal Behaviour, 8, 719-724.

105. Perfetti, Charles and Susan Goldman (1976). Discourse memory and reading comprehension, Journal of Verbal Learning and Learning Behaviour, 15, 33-42.

106. Perfetti, Charles and T. Hogaboam (1975). Relationship between simple word decoding and reading comprehension skill. Journal of Educational Psychology, 67, 461-469.

107. Perren, G.E. and J.L. Trim, ed. (1971). Applications in linguistics. Cambridge : Cambridge University Press.

108. Pilliner, Sheila (1965). A comparison of two English language proficiency test batteries. Unpublished dissertation.

Dept. of Applied Linguistics, University of Edinburgh.

109. Pyrczak, Fred (1975). A responsive note on measures of the passage dependence of reading comprehension test items, Reading Research Quarterly, XI, 112-117.

110. Rankin, Earl (1970). The cloze procedure - its validity and utility. In Farr (1970). Printed earlier in Caussey and Eller (1959 : 131-144). Starting and improving college reading programmes. Fort Worth,: Texas Christian University Press.

111. Reed, David W. (1970a). Linguistic forms and the process of reading. In Levin and Williams (1970 : 19-29).

112. Reed, David W. (1970b). A theory of language, speech and writing. In Singer and Ruddell (1970 : 219-238).

113. Rivers, Wilga (1968). Foreign language skills. Chicago : University of Chicago Press.

114. Robinson, Francis P. (1940). Speed versus comprehension in reading, Journal of Educational Psychology, 31, 554-558.

115. Robinson, Francis P. and Prudence Hall (1941). Studies of higher level reading abilities, Journal of Educational Psychology, 32, 241-252.

116. Robinson, Helen, ed. (1968). Innovation and change in reading instruction. Chicago : University of Chicago Press.

117. Rothkopf, Ernest Z. (1972). Structural text features and the control of processes in learning from written materials. In Carroll and Freedle (1972 : 315-335).

118. Ruddell, Robert B. (1969). Psycholinguistic implications for a systems of communication model. In Goodman and Fleming (1969 : 61-78).

119. Samuels, Jay S., (1970). Factors which influence the word recognition process of skilled readers. In Schick and May (1970, 277-281).

120. Samuels, Jay S. and Patricia Dahl (1975). Establishing appropriate purpose for reading and its effect on flexibility of reading rate, Journal of Educational Psychology, 67, 38-43.

121. Schallert, Diane L. (1976). Improving memory for prose : the relationship between depth of processing and context, Journal of Verbal Learning and Verbal Behaviour, 15, 621-632.

122. Schick, George and Merrill May, ed. (1970). Reading : process

and pedagogy, Vol. I. Milwaukee : National Reading Conference.

123. Schiffman, H.R. (1972). Some components of sensation and perception for the reading process, Reading Research Quarterly, VII, 588-612.

124. Schlesinger, I. (1968). Sentence structure and the reading process. The Hague : Mouton.

125. Simmons, R.F. (1972) Some semantic structures for representing English meanings. In Carroll and Freedle (1972 : 71-97).

126. Singer, Harry (1970). Theoretical models of reading : Implications for teaching and research. In Singer and Ruddell (1970 : 147-182).

127. Singer, Harry and R. Ruddell, ed. (1970). Models and process of reading. Newark, Delaware : International Reading Association.

128. Smith, Frank (1971), Understanding reading : a psycholinguistic analysis of reading and learning to read. New York : Holt, Rinehart and Winston Inc.

129. Smith, Frank, ed. (1973). Psycholinguistics and reading. New York : Holt, Rinehart and Winston Inc.

130. Spache, George D. (1963). Toward better reading. Champagne, Illinois : Garrard.

131. Spache, George D. (1964). Ways of defining the reading process, Reading in the Elementary School.

132. Steiner, R., Wiener, M. and W. Cromer (1971). Comprehension training and identification for poor and good readers, Journal of Educational Psychology, 62, 506-513.

133. Sticht, Thomas G. (1972). Learning by listening. In Carroll and Freedle (1972 : 285-314).

134. Taylor, Wilson L. (1953). Cloze procedure : a new tool for measuring readability, Journalism Quarterly, 30, 415-433.

135. Taylor, Wilson L. (1957). 'Cloze' readability scores as indices of individual differences in comprehension and aptitude, Journal of Applied Psychology, 41, 19-25.

136. Thomas, Laurie F. A flow diagram technique for describing the 'meaning structure' of a text. Undated mimeograph, 11 pp.

137. Thomas, Laurie F. and Sheila Augstein (1972). An experimental approach to the study of reading as a learning skill, Research in Education, 8, 28-46.

138. Thurstone, Eric and L. Hafner, ed. (1965). The philosophical and sociological bases of reading. Milwaukee : National Reading Conference.

139. Tinker, Miles (1939). Speed versus comprehension in reading as affected by level of difficulty, Journal of Educational Psychology, 30, 81-94.

140. Tinker, Miles (1940). Discussion following Robinson (1940)

141. Tuinman, J. Jaap (1972). Speculations on cloze as a search procedure. In Greene (1972 : 74-84).

142. Tuinman, J. Jaap (1973). Determining the passage dependency of comprehension questions in 5 major tests, Reading Research Quarterly, IX, 206-223.

143. Tuinman, J. Jaap and Gordon Gray (1972). The effect of reducing the redundancy of written messages by deletion of function words, Journal of Psychology, 82, 299-306.

144. Ure, Jean (1971). Lexical density and register differentiation. In Perren and Trim (1971 : 443-452).

145. Venezky, Richard L. (1976). Theoretical and experimental base for teaching reading. The Hague : Mouton. Printed earlier in Sebok (1974). Current Trends in Linguistics, vol. 12.

146. Venezky, Richard L. and Robert C. Calfee (1970). The Reading competency model. In Singer and Ruddell (1970 : 273-291).

147. Weaver, Wendell (1965). Theoretical aspects of the cloze procedure. In Thurstone and Hafner (1965 : 116-132).

148. Weaver, Wendell and Albert J. Kingston (1963). A factor analysis of the cloze procedure and other measures of reading and language ability, Journal of Communication, 13, 252-261.

149. Weber, Rose-Marie (1970). First grader's use of grammatical context in reading. In Levin and Williams (1970 : 147-163).

150. Weinstein, Rhona and M. Sam Rabinovitch (1971). Sentence structure and retention in good and poor readers, Journal of Educational Psychology, 62, 25-30.

151. Whatmough, Joshua (1958). Mathematical linguistics, (followed by discussion), Eighth International

Congress of Linguistic Proceedings, Oslo, 74-91.

152. Widdowson, H.G. (1973). Directions in the teaching of discourse. In Corder and Roulet (1973 : 65-76).

153. Widdowson, H.G. (1975). Stylistics and the teaching of literature. London : Longman.

154. Widdowson, H.G. Approaches to discourse, Med Ziffern (forthcoming) 151-167.

155. Wilson, K. and J.B. Carroll (1954). Applications of entropy measures to problems of sequential structure. In psycholinguistics : A survey of theory and research problems, Supplement to Journal of Abnormal and Social Psychology, 103-112.

156. Woodworth, R.S. and H. Schlosberg (1954). Experimental psychology, New York : Henry Holt Ltd.

157. Zechmeister, Eugene B. and Jack McKillip (1972). Recall of place on page, Journal of Educational Psychology, 63, 446-453.

APPENDICES

APPENDIX ONE

TEST MATERIALS

1.A.1.

EXPERIMENT II 8

Part I : 15 minutes : Reading the Passage

Part II : 15 minutes : Writing a Summary of the Passage

Part I INSTRUCTIONS

Read 'You have 15 minutes to read this passage, after which it will be taken back. You will then be required to write a brief summary of the passage.'

[After 15 minutes the passage is to be collected.
Distribute single sheet of paper to each subject.]

Part II INSTRUCTIONS

Read 'Now write your name at top of the sheet. You have 15 minutes to write a brief summary of the passage you have just read. The summary may be in note form with the main points only. You needn't use complete sentences.'

To the Tutor:

I'd like everything back, the passage too, if possible.

Thanks.

1.A.2

INFORMATION FOR THE TUTOR

The experiment consists of three sections to be done at different times as convenient BUT only in the order stated. Sections I and II are to be done within 24 hours of each other. Detailed instructions for each section are attached to the exercises for the particular section. The order to be adopted is as follows:

SECTION 1 Reading Comprehension 1 Time needed: 45 minutes.
(This part comprises a cloze passage with 101 deletions.)

SECTION 2 Reading Comprehension 2 Time needed: 45 minutes.
(This part comprises a passage followed by 20 multiple-choice questions.)

SECTION 3 Two passages on URBANISATION. Time needed: 60 minutes.
(This part comprises two passages, one short and one long, on the same topic. The order of the two passages varies which is why they are stapled together. These two exercises also use cloze.)

In addition there is a very simple Evaluation Form which I would like all candidates to fill in. It would be of interest to see how they evaluate themselves when they read in English, since this is likely to be in a field that either interests them or is considered necessary for academic purposes.

Thank you very much for your co-operation.

I would like ALL the material to be returned to me at the following address,

MISS WILLY ENGINEER

DEPARTMENT OF APPLIED LINGUISTICS

UNIVERSITY OF EDINBURGH

14 BUCCLEUCH PLACE

EDINBURGH EH8 9LH

1.A.3.

INSTRUCTIONS TO THE TUTOR: SECTION 1

The following is the text of the instructions for this section. It is to be read out to the class at the beginning.

SECTION 1 Time needed: 45 minutes

Read 'This is an exercise in reading comprehension. It consists of a long passage with blank spaces. You are required to fill in the blank spaces using only one word in each space. You will have to guess what the missing word is likely to be, taking into account the language and meaning of the passage. Work as quickly as you can. If you are not sure and cannot make a guess, move on to the next blank space. Don't waste time. You may be able to fill in the missing word later when you have read more of the passage.

If you do not understand what you have to do, please ask now. You will not be allowed to ask any questions once the exercise has started.

Now write your name and date at the top. You have 45 minutes for this exercise.

Are there any questions?'

(Pause. Please repeat or clarify if necessary. If there are no questions.)

'Please begin now.'

1.A.4.

INSTRUCTIONS TO THE TUTOR: SECTION 2

This is the text of the instructions for this section. It is to be read out to the class at the beginning.

SECTION 2 Time needed: 45 minutes

Read 'This is an exercise in reading comprehension. It consists of a long passage, the same one you did earlier, but this time it is complete. At the end of the passage there are 20* questions you must answer. There are four possible answers to each question, and you are to choose the one option you think is the best answer, and put a tick against it.

Write your name at the top of the reading passage and at the top of the question section. You have 45 minutes for this exercise.

If you do not understand what you have to do, please ask now. You will not be allowed to ask any questions once this exercise has started.

Are there any questions?'

(Pause. Please repeat or clarify if necessary. If there are no questions,)

'Please begin now.'

PLEASE COLLECT ALL PAPERS, INCLUDING THE READING PASSAGE.

THANK YOU.

Note: * This was changed to 25 for version 3.

1.A.5.

INSTRUCTIONS TO THE TUTOR: SECTION 3

This is the text of the instructions for this section. It is to be read out to the class at the beginning.

SECTION 3 Time needed: 60 minutes

Read 'This is an exercise in reading comprehension. It consists of TWO passages, of different lengths, on the same TOPIC. You must do them in the order they are given because the first one will help you to do the second one better.

It is like the first exercise you did earlier. It consists of blank spaces which you are required to fill. Remember only one word is to be used in each space. You will have to guess what the missing word is likely to be, taking into account the language and meaning of the passage. Work as quickly as you can. If you are not sure and cannot guess for a particular blank space, move on to the next one. Don't waste time. You may be able to fill in the missing word later when you have read more of the passage.
YOU ARE REQUESTED TO TAKE 15 minutes FOR THE SHORT PASSAGE AND 45 minutes FOR THE LONG PASSAGE. YOU HAVE ONE HOUR FOR THE TWO PASSAGES.

Now write your name at the top of both passages.
If you do not understand what you have to do, please ask now. You will not be allowed to ask any questions once the exercise has begun.

Are there any questions?'

(Pause. Please repeat or clarify if necessary. It is important for them to realise that they must do the exercise in the order in which they are stapled together. Some students will have the short passage first, while the others will have the long passage to begin with. Apart from this ordering, both sets are identical. If all is clear,)

'Please begin now. Remember 15 minutes for the short passage, and 45 minutes for the long one.'

- END OF INSTRUCTIONS -

To the Tutor:

If possible, could you keep an eye on the time? Perhaps a call at the end of 15 minutes, and one at the end of 45 minutes would help. THANKS.

1.A.6.

EVALUATION FORM

1. NAME_____ DATE_____

2. FIRST LANGUAGE_____

3. NUMBER OF YEARS YOU HAVE STUDIED ENGLISH_____

4. HOW WELL DO YOU READ ENGLISH? (Tick (✓) the box that gives the best answer in your case.)

WHEN I READ ENGLISH, I USUALLY UNDERSTAND

- ☐ VERY LITTLE.
- ☐ LESS THAN FIFTY PERCENT.
- ☐ ABOUT FIFTY PERCENT.
- ☐ MORE THAN FIFTY PERCENT.
- ☐ ALMOST EVERYTHING.

5. I AM STUDYING ENGLISH BECAUSE_____

_____

1.B.1.1.

Passage 1.

NAME_____

DATE_____

READING COMPREHENSION - 1

Complete the following passage by filling in the blank spaces as you think best. Only one word per blank space should be used. You have 45 minutes for this exercise.

The earliest known civilisations like those of Syria, Mesopotamia, and the valley of the Indus, occupied an area roughly coinciding with the distribution of the wild ancestors of wheat and barley, and the civilisations of Europe and Western Asia may be said to be founded on these two plant species.

The domestication of animals for food began at about (1)_____ same time as cereal cultivation. Cattle, sheep, goats and (2)_____ were used earlier while the use of the horse, (3)_____, ox and camel for transport and traction came later. (4)_____ early stages of settled agriculture, with both cultivated cereals (5)_____ domestic herds, have been most studied at the ancient (6)_____ of the Faiyum and Merimde in Egypt, and Sumer (7)_____ Mesopotamia. It is at the latter that the use (8)_____ sheep's wool for clothing is first recorded, probably 4000 (9)_____ before the present era. Dairy farming also seems to (10)_____ begun in Mesopotamia, perhaps 1000 years later.

But wherever (11)_____ was adopted, agriculture had the most profound effects on (12)_____ life. Its primary

function was to ensure a greater (13)_____ more reliable supply of food, but in doing so (14)_____ made possible a vast number of new techniques and (15)_____ which, in their combined effects, transformed man from a (16)_____ creature into the dominant species that he is today. (17)_____ earliest agricultural communities were probably ten to twenty times (18)_____ large as the nomadic groups of earlier primitive times. (19)_____ the two or three hundred persons, specialisation of labour (20)_____ possible, and this in turn gave rise to more (21)_____ and more skilled operations. Among the earliest trades so (22)_____ were those of the weaver, the potter, and the (23)_____ polisher. There was at this stage, however, no major (24)_____ advance, apart from the development of agriculture and stock-breeding. (25)_____ possibility of great technical developments depended, in the first (26)_____, on the acquisition of sufficient agricultural skill to produce (27)_____ surplus of food from the soil. This made possible (28)_____ existance of people, such as skilled flint workers and (29)_____ specialists, who produced no food for themselves.

An important (30)_____ to farming technique was made by the traction plough: (31)_____ plough drawn by oxen. This greatly increased the arable (32)_____ that one man could easily cultivate. Hence the use (33)_____ domestic animals as a source of power was in (34)_____ of fundamental importance. Another technique of the greatest value (35)_____ that of irrigation, which was developed on a large (36)_____

in Egypt and Mesopotamia. The immediate consequence of these (37)_____ is illustrated in the Faiyum: in this prosperous community (38)_____ between 6000 and 5000 B.C. there were two groups (39)_____ silos, numbering sixty-seven and ninety-eight respectively; each held, on (40)_____ average, about 400 kg. of grain, which was probably (41)_____ yield of two or three acres. This figure, then (42)_____ great achievement, is low by modern standards. In the (43)_____'s in England and Wales an average of about 1350 (44)_____ of wheat was produced per acre.

And so the (45)_____ of better farming techniques created conditions wherein man could (46)_____ those activities we associate with civilisation. Civilisation has been (47)_____ the culture of cities. Cities are themselves large gatherings (48)_____ men not engaged on food production. They appear first (49)_____ the valleys of the Nile in Egypt, the Tigris (50)_____ Euphrates in Mesopotamia, the Indus in northern India, and (51)_____ simultaneously the Yangtse-kiang in China. A further, separate civilisation (52)_____ in Central and South America. We have seen that (53)_____ important function of a food surplus is to support (54)_____ in crafts and industries other than agriculture and stock-breeding. (55)_____ surplus, however, does more than that. From the beginning (56)_____ was also a basis for trade. Certain important commodities (57)_____ as flint, and later copper, tin and iron that (58)_____ found only in a few places, were mined and (59)_____ long distances in exchange for food. They were brought (60)_____

in the crude state or in the form of (61)_____ products such as axes, vessels, shields or ornaments; and (62)_____ consumer goods such as pottery and trinkets entered into (63)_____ trade of early civilisation. As human communities became more (64)_____, they came to need a central administration, a legal (65)_____, and machinery for keeping order and enforcing the law. (66)_____ were the centres, therefore, not only of craftsmanship and (67)_____, but also of administration, and the food surplus had (68)_____ support a ruling class of priests and kings who (69)_____ taxes or tribute in the form of food. This (70)_____ the origin of the state, the apparatus of government (71)_____ control by a small ruling class.

The history of (72)_____ civilisations illustrates very clearly the way in which the (73)_____ of science is related to human needs. Between five (74)_____ six thousand years ago the wheel was invented and (75)_____ both for transport and in the making of pottery. (76)_____ trade led to the development of the sailing ship, (77)_____ this in turn led to the establishing of navigational (78)_____. Navigation requires a knowledge of the stars, and so (79)_____ science of astronomy was encouraged. For agriculture one needed (80)_____ plan the sowing of crops according to the seasons, (81)_____ so a calendar was invented. The Egyptians had a (82)_____ of 365 days and no leap year, and this (83)_____ was probably first used in 4242 B.C. Knowledge of (84)_____ and the working of the calendar was no doubt (85)_____ privilege of the priest-kings, and since this

knowledge was (86)_____ source of power, we can guess that they guarded (87)_____ jealously; they provide an early example of the restriction (88)_____ learning to a particular class.

With the growth of (89)_____, various technical problems arose which had to be solved. (90)_____ of them was the need to supply clean water (91)_____ ensure the disposal of waste. At the city of (92)_____, on the River Indus, there was a well-organized system (93)_____ aqueducts and drains, and there is evidence of town-planning (94)_____ the city government.

Civilisation not only creates material wealth (95)_____ an elaborate technology, but by making leisure possible it (96)_____ to the development of the arts. The artists of (97)_____ primitive times took advantage of a good supply of (98)_____ to use their leisure to paint and draw on (99)_____ walls and on bone. The artists of civilisation could (100)_____ express themselves through architecture as well, and in Egypt (101)_____ musical scale was invented. In this way on the basis of a successful agriculture, the main features of all civilisations up to 200 years ago were developed nearly 5000 years before the industrial revolution in Europe.

1.B.1.2.

PASSAGE 1 : ACCEPTABLE WORD RESPONSES

Item No.	Exact Word	Acceptable Words
1	the	
2	pigs	hens, poultry, chickens
3	ass	donkey, mule, elephant, buffalo
4	the	these, those
5	and	
6	settlements	sites, settlement
7	in	
8	of	
9	years	
10	have	
11	it	this
12	human	man's, daily, everyday, rural, one's, everybody's
13	and	
14	it	this
15	activities	skills, crafts, devices, processes
16	rare	primitive, dominated, helpless, vulnerable, weak
17	the	these
18	as	
19	among	amongst
20	became	was
21	complex	technical, specialised, specific, efficient, varied, complicated, effective, planned, organized, elaborate, sophisticated

Item No.	Exact Word	Acceptable Words
22	established	developed, created, acquired
23	flint	stone, metal
24	technical	economic, technological, scientific
25	the	any
26	place	instance
27	a	enough, some, sufficient
28	the	
29	other	similar, tool, various
30	contribution	addition
31	the	a
32	land	area
33	of	
34	fact	itself
35	was	is
36	scale	
37	advances	techniques, improvements, developments, innovations, practices, experiments
38	of	founded, established, settled, roughly, probably
39	of	possessing, owning
40	the	an
41	the	
42	a	
43	1950's	
44	kg	
45	practice	acquisition, evolution, development, use, application, implementation, introduction

Item No.	Exact Word	Acceptable Words
46	invent	start, practice, develop, pursue, promote, initiate
47	called	largely, especially
48	of	
49	in	along
50	and	
51	perhaps	almost, also, maybe, probably
52	arose	appeared, emerged, developed, flourished, began, rose
53	an	the, one
54	workers	men, workmen, people, progress, growth, activities, development, developments
55	the	a, this
56	it	food, this, there
57	such	
58	are	were
59	transported	carried, taken, sent
60	either	
61	finished	manufactured, ready-made, various, man-made
62	other	several, certain, some, also, thus
63	the	
64	complex	sophisticated, advanced, civilised
65	system	code, framework, organization, structure
66	cities	they, these, towns
67	trade	industry
68	to	

Item No.	Exact Word	Acceptable Words
69	took	collected, levied, imposed, introduced, exacted, demanded, raised, claimed
70	was	
71	and	or
72	ancient	early, older, these, many, earlier, several, old
73	development	progress, advancement, advance, evolution, role
74	and	to
75	used	
76	oversea	overseas, more, early, sea, increasing, maritime, expanding, growing, coastal
77	and	
78	methods	science, routes, skills, techniques, aids, practices, systems, studies
79	the	
80	to	
81	and	
82	year	calendar
83	calendar	one, system
84	astronomy	stars, science
85	a	the
86	a	their
87	it	them
88	of	
89	cities	towns, population, civilisation, science, communities
90	one	

	Item No.	Exact Word	Acceptable Words
	91	and	
*	92	Mohenjo-daro	
	93	of	
	94	by	in
	95	and	with, for, by
	96	leads	contributes, contributed, adds, led
	97	earlier	these, more, very
	98	meat	food
	99	cave	stone, both, rock, mud, their, stony
	100	also	moreover, now, happily, certainly, indeed
	101	a	the

* Note: Item 92 not included in calculations.

1.B.1.3

Name _____

Date _____

READING COMPREHENSION - 2

Read the following passage carefully and then answer the questions given at the end. You have 45 minutes for this exercise.

The earliest known civilisations like those of Syria, Mesopotamia, and the valley of the Indus, occupied an area roughly coinciding with the distribution of the wild ancestors of wheat and barley, and the civilisations of Europe and Western Asia may be said to be founded on these two plant species.

The domestication of animals for food began at about the same time as cereal cultivation. Cattle, sheep, goats and pigs were used earlier while the use of the horse, ass, ox and camel for transport and traction came later. The early stages of settled agriculture, with both cultivated cereals and domestic herds, have been most studied at the ancient settlements of the Faiyum and Merimde in Egypt, and Sumer in Mesopotamia. It is at the latter that the use of sheep's wool for clothing is first recorded, probably 4000 years before the present era. Dairy farming also seems to have begun in Mesopotamia, perhaps 1000 years later.

But wherever it was adopted, agriculture had the most profound effects on human life. Its primary function was to ensure a greater and more reliable supply of food, but in doing so it made possible a vast number of new techniques and activities which, in their combined effects, transformed man from a rare creature into the dominant species that he is today. The earliest agricultural communities were probably ten to twenty times as large as the nomadic groups of earlier primitive times. Among the two or three hundred persons, specialisation of labour became possible, and this in turn gave rise to more complex and more skilled operations. Among the earliest trades so established were those of the weaver, the potter, and the flint-polisher. There was at this stage, however, no major technical advance, apart from the development of agriculture and stock-breeding. The possibility of great technical developments depended, in the first place, on the acquisition of sufficient agricultural skill to produce a surplus of food from the soil. This made possible the existance of people, such as skilled flint workers and other specialists, who produced no food for themselves.

An important contribution to farming technique was made by the traction plough - the plough drawn by oxen. This greatly increased the arable land that one man could easily cultivate. Hence the use of domestic animals

as a source of power was in fact of fundamental importance. Another technique of the greatest value was that of irrigation, which was developed on a large scale in Egypt and Mesopotamia. The immediate consequence of these advances is illustrated in the Faiyum: in this prosperous community of between 6000 and 5000 B.C. there were two groups of silos, numbering sixty-seven and ninety-eight respectively; each held, on the average, about 400 kg. of grain, which was probably the yield of two or three acres. This figure, then a great achievement, is low by modern standards. In the 1950's in England and Wales an average of about 1350 kg. of wheat was produced per acre.

And so the practice of better farming techniques created conditions wherein man could invent those activities we associate with civilisation. Civilisation has been called the culture of cities. Cities are themselves large gatherings of men not engaged on food production. They appear first in the valleys of the Nile in Egypt, the Tigris and Euphrates in Mesopotamia, the Indus in northern India, and perhaps simultaneously the Yangtse-kiang in China. A further, separate civilisation arose in Central and South America. We have seen that an important function of a food surplus is to support workers in crafts and industries other than agriculture and stock-breeding. The surplus, however, does more than that. From the beginning it was also a basis for trade. Certain important commodities such as flint, and later copper, tin and iron that are found only on a few places, were mined and transported long distances in exchange for food. They were brought either in the crude state or in the form of finished products such as axes, vessels, shields or ornaments; and other consumer goods such as pottery and trinkets entered into the trade of early civilisation. As human communities became more complex, they came to need a central administration, a legal system, and machinery for keeping order and enforcing the law. Cities were the centres, therefore, not only of craftsmanship and trade, but also of administration, and the food surplus had to support a ruling class of priests and kings who took taxes or tribute in the form of food. This was the origin of the state, the apparatus of government and control by a small ruling class.

The history of ancient civilisations illustrates very clearly the way in which the development of science is related to human needs. Between five and six thousand years ago the wheel was invented and used both for transport and in the making of pottery. Oversea trade led to the development of the sailing ship, and this in turn led to the establishing of navigational methods. Navigation requires a knowledge of the stars, and so the science of astronomy was encouraged. For

agriculture one needed to plan the sowing of crops according to the seasons, and so a calendar was invented. The Egyptians had a year of 365 days and no leap year, and this calendar was probably first used in 4242 B.C. Knowledge of astronomy and the working of the calendar was no doubt a privilege of the priest-kings, and since this knowledge was a source of power, we can guess that they guarded it jealously; they provide an early example of the restriction of learning to a particular class.

With the growth of cities, various technical problems arose which had to be solved. One of them was the need to supply clean water and ensure the disposal of waste. At the city of Mohenjo-daro, on the River Indus, there was a well-organized system of aqueducts and drains, and there is evidence of town-planning by the city government.

Civilisation not only creates material wealth and an elaborate technology, but by making leisure possible it leads to the development of the arts. The artists of earlier primitive times took advantage of a good supply of meat to use their leisure to paint and draw on cave walls and on bone. The artists of civilisation could also express themselves through architecture as well, and in Egypt a musical scale was invented. In this way on the basis of a successful agriculture, the main features of all civilisations up to 200 years ago were developed nearly 5000 years before the industrial revolution in Europe.

1.B.1.4.

MULTIPLE-CHOICE: TRY-OUT VERSION

Answer the following questions choosing the one option that best fits with the information given in the passage. Put a tick (✓) in the corresponding box.

1. Dairy farming seems to have begun in Mesopotamia. What else is also first recorded there?
   (a) a drainage system
   (b) the use of sheep's wool for clothing
   (c) an irrigation system on a large scale
   (d) silos for storing grain
2. What does 'it' in line 16 refer to?
   (a) agriculture
   (b) sheep's wool for clothing
   (c) domestication of animals
   (d) dairy farming
3. Trade first became possible when
   (a) sailing ships were invented
   (b) consumer goods such as pottery and trinkets were made
   (c) human communities became more complex
   (d) food could be exchanged for other commodities
4. The earliest known civilisations roughly coincided with areas of cultivation because
   (a) civilisation is the culture of cities
   (b) all the farming population lived there
   (c) with surplus food other professions could develop
   (d) the primary function of agriculture was to provide a

greater supply of food

5. How does the author define cities?

   (a) large gatherings of men engaged in agriculture

   (b) a place where artists could express themselves through architecture

   (c) large gatherings of men engaged in occupations other than agriculture

   (d) a place where various technical problems needed to be solved

6. Ancient civilisations grew up where

   (a) animals could easily be hunted for food

   (b) a wild variety of wheat and barley were found

   (c) domestication of animals was possible

   (d) nomadic groups had lived in earlier times

7. What does 'This' in line 67 refer to?

   (a) centres of administration controlled by a few

   (b) a central administration, a legal system and measures to enforce it

   (c) a ruling class who taxed the people

   (d) a machinery for keeping law and order

8. What does 'this' in line 30 refer to?

   (a) the possibility of great technical advance

   (b) great technical advance

   (c) acquisition of sufficient agricultural skill

   (d) surplus food

9. Specialisation of labour became possible as a direct result of

   (a) the domestication of animals

   (b) cereal cultivation by farmers

(c) the use of animals for food and transport

(d) the availability of food for their support

10. The civilisation in Egypt is best known for

(a) the use of dairy farming

(b) the building of a good drainage system

(c) the development of settled agriculture

(d) a system of aqueducts

11. The early nomadic groups had an average population of

(a) ten to twenty

(b) two to three hundred

(c) ten to thirty

(d) two to three thousand

12. Technical development depended in the beginning on

(a) a skilled tradesman like the flint worker

(b) agricultural know-how

(c) the traction plough

(d) animals as a source of power

13. Who does 'a particular class' in lines 83 and 84 refer to ?

(a) priest-kings

(b) navigators

(c) farmers

(d) astronomers

14. Aqueducts and drains are given as examples of

(a) technical problems created by the growth of cities

(b) elaborate technology created by civilisation

(c) scientific developments related to human needs

(d) civilisation itself

15. The information contained in the passage is based on

(a) historical facts

(b) scientific discoveries

(c) the author's opinions

(d) calculated guesswork

16. 'The earliest known civilisations occupied an area roughly coinciding with the distribution of the wild ancestors of wheat and barley.' The word 'roughly' here means

(a) exactly

(b) generally

(c) slightly

(d) not

17. What does 'they' in line 49 refer to?

(a) civilisations

(b) large gatherings of men

(c) cities

(d) better farming techniques

18. In line 53 read the sentence beginning with the words, 'We have seen that ...' Where has this information been given?

(a) paragraph 1

(b) paragraph 2

(c) paragraph 3

(d) paragraph 4

19. According to the author, what is the probable chronological order of the following events:

1. use of domestic animals as a source of power
2. use of domestic animals for food
3. establishment of various trades
4. development of oversea trade
5. acquisition of agricultural skills to produce

extra food

(a) 1,2,3,4,5

(b) 2,1,5,3,4

(c) 4,1,2,5,3

(d) 5,2,1,4,3

20. The best summary of the passage is contained in the following sentence.

(a) This was the origin of the state, the apparatus of government and control by a small ruling class.

(b) The history of ancient civilisation illustrates very clearly the way in which the development of science is related to human needs.

(c) The practice of better farming techniques created conditions wherein man could invent those articles we associate with civilisation.

(d) Civilisation not only creates material wealth and an elaborate technology but by making leisure possible it leads to the development of the arts.

1.B.1.5

MULTIPLE-CHOICE QUESTIONS REFLECTING TAXONOMY OF ABILITIES

PRIMARY ABILITY	LOCALISED DISTRACTORS					NON-LOCALISED DISTRACTORS				
FACTUAL							1		6 16	
REFERENTIAL	3		5 10			7				
INFERENTIAL		9			4 8	2 12			11	19
EVALUATIVE			15			14 18		13		17 20
OTHER SKILLS										
SECONDARY ABILITY TESTED ⇒	FACTUAL	REFERENTIAL	INFERENTIAL	EVALUATIVE	OTHER SKILLS	FACTUAL	REFERENTIAL	INFERENTIAL	EVALUATIVE	OTHER SKILLS

VERSIONS 1 & 2

PRIMARY ABILITY	LOCALISED DISTRACTORS					NON-LOCALISED DISTRACTORS				
FACTUAL							1		16	
REFERENTIAL	3		5 10			7				
INFERENTIAL		9			4 8	2 12			11	19
EVALUATIVE			15			14 18		13	24	6,17, 20,21, 22,23,25
OTHER SKILLS										

SECONDARY ABILITY TESTED	FACTUAL	REFERENTIAL	INFERENTIAL	EVALUATIVE	OTHER SKILLS	FACTUAL	REFERENTIAL	INFERENTIAL	EVALUATIVE	OTHER SKILLS

VERSION 3

ITEMS CONCERNED WITH TEXTUAL STRUCTURE : 6,19,21,22,23,25

1.B.1.6.

MULTIPLE-CHOICE : VERSION 1

NAME_____

DATE_____

Answer the following questions choosing the one option that best fits with the information given in the passage. Put a tick (✓) in the corresponding box.

1. Dairy farming seems to have begun in Mesopotamia.
   What else is also first recorded there?
   (a) a drainage system
   (b) the use of sheep's wool for clothing
   (c) an irrigation system on a large scale
   (d) silos for storing grain

2. Specialisation of labour became possible as a direct result of
   (a) the domestication of animals
   (b) cereal cultivation by farmers
   (c) the use of animals for traction and transport
   (d) the availability of surplus food

3. What does 'it' in line 16 refer to?
   (a) agriculture
   (b) sheep's wool for clothing
   (c) domestication of animals
   (d) dairy farming

4. How does the author define cities?
   (a) a place where various technical problems needed to be solved

(b) a place where artists could express themselves through architecture

(c) large gatherings of men engaged in agriculture

(d) large gatherings of men engaged in occupations other than agriculture

5. What does 'this' in line 30 refer to?

(a) the possibility of great technical advance

(b) the existence of people

(c) acquisition of skill to produce surplus food

(d) acquisition of sufficient agricultural skill

6. Trade first became possible when

(a) surplus food could be exchanged for other commodities

(b) sailing ships were developed

(c) commodities like flint, copper, tin and iron were found

(d) consumer goods such as pottery and trinkets were made

7. In line 53, read the sentence beginning with the words, 'We have seen that ...' Where has this information been given before?

(a) paragraph 1

(b) paragraph 2

(c) paragraph 3

(d) paragraph 4

8. The early nomadic groups had an average population of

(a) ten to twenty

(b) ten to thirty

(c) two to three hundred

(d) two to three thousand

9. 'The earliest known civilisations ... occupied an area <u>roughly</u>

coinciding with the distribution of the wild ancestors of wheat and barley ...' (See paragraph 1). The word roughly here means

(a) exactly

(b) approximately

(c) slightly

(d) not

10. Who does 'a particular class' in lines 83 and 84 refer to?

(a) priest-kings

(b) navigators

(c) farmers

(d) astronomers

11. The earliest known civilisations coincided roughly with areas of cultivation because

(a) civilisation is the culture of cities

(b) all the farming population lived there

(c) with surplus food other professions could develop

(d) the primary function of agriculture was to provide a greater supply of food

12. Aqueducts and drains are given as examples of

(a) scientific developments related to human needs

(b) civilisation itself

(c) an elaborate technology created by civilisation

(d) technical problems created by the growth of cities

13. According to the passage, which one of the following pieces of information is based on actual fact?

(a) storage facilities for grain have existed for more than 6000 years (lines 39 to 41)

(b) dairy farming began in 3000 B.C. (lines 14 to 15)

(c) only the Egyptian priest-kings could use the

calendar in 4242 B.C. (lines 78 to 80)

(d) the average yield in ancient Egypt was 130 to 200 kg. per acre (lines 41 to 43)

14. Technical development depended in the beginning on

(a) a skilled tradesman like the flint-worker

(b) the traction plough

(c) agricultural know-how

(d) animals as a source of power

15. Four of the following conditions were needed for the origin of the state (line 67). Which four are these?

1. a food surplus
2. a central administration
3. a legal system
4. a machinery for enforcing law and order
5. control by a small group
6. growth of cities

(a) 1,2,3,4

(b) 2,3,4,5

(c) 3,4,5,6

(d) 1,2,5,6

16. Ancient civilisations grew up where

(a) animals could easily be hunted for food

(b) a wild variety of wheat and barley were found

(c) domestication of animals was possible

(d) nomadic groups had lived in earlier times

17. What is the central theme of the passage?

(a) the development of technology

(b) the origin of the state

(c) the growth of cities

(d) the development of civilisation

18. The civilisation in Egypt is best known for

(a) the use of dairy farming

(b) the building of a good drainage system

(c) the development of settled agriculture

(d) a system of aqueducts

19. According to the author, what is the probable chronological order of the following?

1. use of domestic animals for food
2. use of animals for ploughing
3. town-planning
4. painting on cave-walls
5. invention of the wheel

(a) 4,1,2,5,3

(b) 2,1,4,5,3

(c) 1,2,5,4,3

(d) 1,2,3,4,5

20. The best summary of the passage is contained in the following sentence.

(a) Wherever it was adopted, agriculture had the most profound effects on human life.

(b) The practice of better farming techniques created conditions wherein man could invent those articles we associate with civilisation.

(c) The history of ancient civilisation illustrates very clearly the way in which the development of science is related to human needs.

(d) On the basis of a successful agriculture the main features of all civilisations were developed nearly 5000 years ago.

1.B.1.7.

MULTIPLE-CHOICE : VERSION 2

NAME_____

DATE_____

Answer the following questions choosing the one option that best fits with the information given in the passage. Put a tick (✓) against the corresponding letter: (a), (b), (c), or (d).

1. Dairy farming seems to have begun in Mesopotamia. What else is also first recorded there?

   (a) the traction plough

   (b) the use of sheep's wool for clothing

   (c) an irrigation system on a large scale

   (d) silos for storing grain

2. Specialisation of labour became possible as a direct result of

   (a) the domestication of animals

   (b) cereal cultivation by farmers

   (c) the use of animals for traction and transport

   (d) the availability of surplus food

3. What does 'it' in line 16 refer to?

   (a) agriculture

   (b) sheep's wool for clothing

   (c) domestication of animals

   (d) dairy farming

4. Which one of the following would the author exclude from his definition of 'cities'?

   (a) centres of administration

(b) centres of craftmanship and trade

(c) large gatherings of men engaged in agriculture

(d) large gatherings of men engaged in occupations other than agriculture

5. What does 'this' in line 30 refer to?

(a) the possibility of great technical advance

(b) the existence of skilled workers

(c) acquisition of skill to produce surplus food

(d) acquisition of sufficient agricultural skill

6. Trade first became possible when

(a) surplus food could be exchanged for other commodities

(b) the wheel was invented

(c) commodities like flint, copper, tin and iron were found

(d) consumer goods such as pottery and trinkets were made

7. In line 53 read the sentence beginning with the words, 'We have seen that...' Where has this information been given before?

(a) paragraph 1

(b) paragraph 2

(c) paragraph 3

(d) paragraph 4

8. The early nomadic groups had an average population of

(a) twenty to thirty

(b) ten to thirty

(c) two to three hundred

(d) two to three thousand

9. 'This figure, then a great achievement, is low by modern standards ' (lines 43 to 44). The word 'then' in the context

means

(a) subsequently

(b) later

(c) at this time

(d) at that time

10. Who does 'a particular class' in lines 83 and 84 refer to?

(a) priest-kings

(b) navigators

(c) Egyptians

(d) astronomers

11. The earliest known civilisations coincided roughly with areas of cultivation because

(a) civilisation is the culture of cities

(b) all the farming population lived there

(c) with surplus food other professions could develop

(d) the primary function of agriculture was to provide a greater supply of food

12. Aqueducts and drains are given as specific examples of

(a) scientific developments related to human needs

(b) the growth of technical problems created by cities

(c) an elaborate technology created by material wealth

(d) solutions to technical problems created by the growth of cities

13. According to the passage, which <u>one</u> of the following pieces of information is based on actual fact?

(a) storage facilities for grain have existed for more than 6000 years (lines 39 to 41)

(b) dairy farming began in 3000 B.C. (lines 14 to 15)

(c) only the Egyptian priest-kings could use the calendar in 4242 B.C. (lines 78 to 80)

(d) the average yield in ancient Egypt was 130 to 200 kg. per acre (lines 41 to 43)

14. Technical development depended in the beginning on

(a) a skilled tradesman like the flint-worker

(b) the traction plough

(c) agricultural know-how

(d) animals as a source of power

15. Four of the following conditions were needed for the origin of the state (see line 67). Which four are these?

1. complex human communities
2. a food surplus
3. a central administration
4. a machinery for enforcing law and order
5. control by a small group
6. a growth of cities

(a) 1,2,3,4

(b) 2,3,4,5

(c) 3,4,5,6

(d) 1,2,5,6

16. Ancient civilisations grew up where

(a) animals could easily be hunted for food

(b) a wild variety of wheat and barley were found

(c) domestication of animals was possible

(d) nomadic groups had lived in earlier times

17. What is the central theme of the passage?

(a) the history of science

(b) the origin of the state

(c) the development of technology

(d) the development of civilisation

18. The civilisation in Egypt is best known for

(a) the invention of a musical scale

(b) the domestication of animals

(c) the development of settled agriculture

(d) the building of a good drainage system

19. According to the author, what is the probable chronological order of the following?

1. use of animals for ploughing
2. specialisation of labour
3. town-planning
4. use of domestic animals for food
5. invention of the sailing ship

(a) 4,1,2,5,3

(b) 2,1,4,5,3

(c) 1,2,5,4,3

(d) 1,2,3,4,5

20. The best summary of the passage is contained in the following sentence.

(a) The practice of better farming techniques created conditions wherein man could invent those articles we associate with civilisation.

(b) The growth of cities therefore led to the origin of the state, the apparatus of government and control by a small class.

(c) The history of science illustrates very clearly the

way in which its development was related to human needs.

(d) On the basis of a successful agriculture the main features of all civilisations were developed nearly 5000 years ago.

1.B.1.8.

MULTIPLE-CHOICE : VERSION 3

NAME_____

DATE_____

Answer the following questions choosing the one option that best fits with the information given in the passage. Put a tick (✓) against the corresponding letter: (a), (b), (c), or (d).

1. Dairy farming seems to have begun in Mesopotamia. What else is also first recorded there?

   (a) the traction plough

   (b) the use of sheep's wool for clothing

   (c) an irrigation system on a large scale

   (d) silos for storing grain

2. Specialisation of labour became possible as a direct result of

   (a) the domestication of animals

   (b) cereal cultivation by farmers

   (c) the use of animals for traction and transport

   (d) the availability of surplus food

3. What does 'it' in line 16 refer to?

   (a) agriculture

   (b) sheep's wool for clothing

   (c) domestication of animals

   (d) dairy farming

4. Which one of the following would the author exclude from his definition of 'cities'?

   (a) centres of administration

(b) centres of craftsmanship and trade

(c) large gatherings of men engaged in agriculture

(d) large gatherings of men engaged in occupations other than agriculture

5. What does 'this' in line 30 refer to?

(a) the possibility of great technical advance

(b) the existence of skilled workers

(c) the acquisition of skill to produce surplus food

(d) the acquisition of sufficient agricultural skill

6. Read the sentence beginning 'And so the practice of better... ' (paragraph 5, lines 46 to 47). What is its function?

(a) to act as link between the first and second halves of the passage

(b) to summarise and conclude the first half of the passage

(c) to introduce the second half of the passage

(d) to provide the key or topic sentence of the new paragraph

7. In line 53, read the sentence beginning with the words, 'We have seen that ...' Where has this information been given before?

(a) paragraph 1

(b) paragraph 2

(c) paragraph 3

(d) paragraph 4

8. What does 'present era' in line 13 mean?

(a) the time when the passage was written

(b) A.D.

(c) the twentieth century in general

(d) the meaning remains unclear from the passage

9. 'This figure, then a great achievement, is low by modern standards ' (lines 43 to 44). The word 'then' in the context means

(a) subsequently

(b) later

(c) at this time

(d) at that time

10. Who does 'a particular class' in lines 83 and 84 refer to?

(a) priest-kings

(b) navigators

(c) Egyptians

(d) astronomers

11. The earliest known civilisations coincided roughly with areas of cultivation because

(a) civilisation is the culture of cities

(b) all the farming population lived there

(c) with surplus food other professions could develop

(d) the primary function of agriculture was to provide a greater supply of food

12. Aqueducts and drains are given as specific examples of

(a) scientific developments related to human needs

(b) the growth of technical problems created by cities

(c) an elaborate technology created by material wealth

(d) solutions to technical problems created by the growth of cities

13. In the passage some of the information is based on actual fact,

and some on assumption. According to the passage, only <u>one</u> of the following statements is based on fact. Which is it?

(a) storage facilities for grain have existed for more than 6000 years (lines 39 to 41)

(b) dairy farming began in 3000 B.C. (lines 14 to 15)

(c) the Egyptian calendar was first invented in 4242 B.C. (lines 78 to 80)

(d) the average yield in ancient Egypt was about 200 kg. per acre (lines 41 to 43)

14. Technical development depended in the beginning on

(a) a skilled tradesman like the flint-worker

(b) the traction plough

(c) agricultural know-how

(d) animals as a source of power

15. <u>Four</u> of the following conditions were needed for the origin of the state (see lines 67 to 69). Which four are these?

(a) 1,2,3,4

(b) 2,3,4,5

(c) 3,4,5,6

(d) 1,2,5,6

1. a growth of cities
2. a central administration
3. complex human communities
4. a food surplus
5. control by a small ruling class
6. some form of taxation or tribute

16. Ancient civilisations grew up where

(a) animals could easily be hunted for food

(b) a wild variety of wheat and barley were found

(c) domestication of animals was possible

(d) nomadic groups had lived in earlier times

17. What is the central theme of the passage?

(a) the history of science

(b) the origin of the state

(c) the development of technology

(d) the development of civilisation

18. The civilisation in Egypt is best known for

(a) the invention of a musical scale

(b) the domestication of animals

(c) the development of settled agriculture

(d) the building of a good drainage system

19. According to the author, what is the probable chronological order of the following?

(a) 4,1,2,5,3

(b) 2,1,4,5,3

(c) 1,2,5,4,3

(d) 1,2,3,4,5

1. the invention of the wheel
2. the development of modes of transport
3. the science of astronomy
4. the use of domestic animals for food
5. the development of the sailing ship

20. The <u>best summary</u> of the passage is contained in the following sentence.

(a) The practice of better farming techniques created conditions wherein man could invent those articles

we associate with civilisation.

(b) The growth of cities therefore led to the origin of the state, the apparatus of government and control by a small ruling class.

(c) The history of science illustrates very clearly the way in which its development was related to human needs.

(d) On the basis of a successful agriculture the main features of all civilisations were developed nearly 5000 years ago.

21. The function of the first paragraph is

(a) to introduce the main theme of the passage

(b) as an introduction to the theme which comes later in the passage

(c) to present a fact related to the main theme

(d) to provide background information only

22. What is the function of the second paragraph?

(a) to introduce the main theme along with supporting details

(b) to provide a logical development to the theme introduced in the first paragraph

(c) to present a second but unconnected fact along with supporting details

(d) to present a series of details in support of the main theme introduced in the first paragraph

23. What, according to you, is the probable source of this passage?

(a) a 'text' book for study purposes

(b) a 'popular' book for general information

(c) an 'encyclopaedia' for reference purposes

(d) an 'article' from a newspaper

24. Read paragraph 2, lines 12 to 15. What is the author uncertain about?

(a) the sequence of his facts

(b) the facts themselves

(c) the dates of the facts

(d) all the above, that is, (a), (b) and (c)

25. Why is 'but' used in the opening sentence of the third paragraph?

(a) to contrast the information in the sentence with that in paragraph 2.

(b) for emphasis of the information in the sentence itself

(c) to contradict earlier information

(d) for stylistic reasons only

1.B.2.1.

Passage 2A

NAME_____

DATE_____

WHAT IS LAW?

Read the following passage and complete it by filling in the blank spaces with the most suitable word you can think of. Only one word per blank space should be used. You have 45 minutes for this exercise.

Law has been defined as 'the just interference of the State in the interests and passions of humanity'. It will hardly be seriously disputed that it is (1)_____ for the State to interfere in the affairs of (2)_____ community to a greater or lesser degree. Community life (3)_____ be a sad tangle if there were no laws (4)_____ rules which had to be observed by all.

The (5)_____ of us are, apparently, prepared to abandon a certain (6)_____ of freedom, because we realize 'controls' are essential to (7)_____ reconstruction of the world, in much the same way (8)_____ scaffolding is required in the erection of a new (9)_____. When the building is finished, the scaffolding may be (10)_____, and when the world has been rebuilt we may (11)_____ for the removal of the present restrictive controls. But (12)_____ does not mean that even then we can dispense (13)_____ all our rules. For example, when you are out (14)_____ after dark without a light, and a policeman stops (15)_____, you

may be infuriated at this attack on your (16)_____. When however, you appear in Court a few days (17)_____, charged with the legal offence of cycling after dark (18)_____ a light, you are not charged with this offence (19)_____ the 'Government' is tyrannical. It is the duty of (20)_____ government to protect the community from the antics of (21)_____. It therefore requires cyclists to carry lamps at night (22)_____ it has a greater duty to safeguard the community (23)_____ a whole than to respect the freedom of the (24)_____.

If, in fact, everyone respected the rights of every (25)_____ member of the community, we should not require any (26)_____ or law to deal with or punish crime. As, (27)_____, crime is in our marrow, so to speak, it (28)_____ necessary to define the exact nature of every criminal (29)_____ and prescribe the punishment to be exacted for breaking (30)_____ particular rule.

Many rules which involve legal obligations are (31)_____ by Acts of Parliament, and they are then known (32)_____ Statute Law. Statute-making became a regular practice in the (33)_____ century. Before any rules derived from an Act of (34)_____ have the force of law today, the Government, or (35)_____ of the Members of the House of Lords or (36)_____ House of Commons, must present a Bill setting out (37)_____ proposed rules in exact terms. The Bill must be (38)_____ and approved three times in the House of Commons, (39)_____ also in the House of Lords, before it can (40)_____ the formal assent of the Queen. Thereafter it

becomes (41)_____, and is placed upon the Statute Book. No law (42)_____ be enacted without the direct or indirect approval of (43)_____ House of Commons. As England is a democracy, and (44)_____ Members of Parliament who constitute the government are elected (45)_____ the people, such laws may be said, accordingly, to (46)_____ the will of the people.

Statute Law, however, is (47)_____ to only some of the problems which arise in (48)_____ daily contacts which take place between individual members of (49)_____ community. When, therefore, a dispute is brought before the (50)_____ for decision, and there is no Statute Law applicable, (51)_____ Judge must find his solution to the problem by (52)_____ methods. After he has heard all the facts, he (53)_____ consider what recognized legal principles are to be applied (54)_____ he delivers his judgement. And when a judgement is (55)_____ in the High Court of Justice, the judgement almost (56)_____ becomes part of the law of the land, in (57)_____ same way as a Statute. It is then known (58)_____ described as part of the 'Common Law'. Common Law (59)_____ therefore of an entirely different character from Statute Law, (60)_____ dates back to the Middle Ages when Courts were (61)_____ in the country. From the beginning, cases came for (62)_____ before the Courts which could not be decided by (63)_____ to any law on the Statute Book. For the (64)_____ of convenience, it therefore became the practice when a (65)_____ delivered his judgement for it to become what is (66)_____ a

'precedent', and in this way the Common Law (67)_____ built up on precedent. In any other case which (68)_____ similar facts, the Judge who tried it had to (69)_____ a similar judgement. Today every judge is bound by (70)_____, and only Parliament, by passing of a new Statute, (71)_____ remake the law on the point in question. The (72)_____ 'Common Law' is, accordingly, habitually used by lawyers to (73)_____ a case which is founded on 'precedent' as distinct (74)_____ Statute Law, but the Common Law may never over-ride (75)_____ Law.

What does the average citizen mean when he (76)_____ he is 'going to law'? He means that he (77)_____ to invoke the assistance of the Courts, in order (78)_____ assert what he believes to be his rights. Since (79)_____ is bound by law, everyone is entitled to ask (80)_____ its assistance against a defaulting member of the community.

(81)_____ man goes to law because he believes the laws (82)_____ fairly administered, and subject to all-too-many human imperfections, he (83)_____ not far wrong. His complaint will be listened to (84)_____ care and he will receive a fair trial. If (85)_____ were not so the results would be disastrous. Indeed, (86)_____ would not bring an action - the technical way of (87)'_____ to law' - unless you felt assured of an impartial (88)_____ of your case. You would not be so foolish (89)_____ to waste your time and money, if you thought (90)_____ result could be influenced by bribery and corruption, or (91)_____ you feared that politics would enter into the case.

(92)_____ then do people scoff at the law? There are (93)_____ doubt a variety of reasons, some of them true (94)_____ some due to ignorance. There are cases where the (95)_____ has been abused and a criminal has escaped conviction, (96)_____ where an innocent man has been wrongfully committed. Lawyers (97)_____ not superhuman nor are they immune from the general (98)_____ of mankind. The problem involved in trying to make (99)_____ rules for every member of a community of millions (100)_____ beyond human capacity. Our best efforts are bound to fall short of perfection; all we can do is support and work for a strong and impartial judicial system as the first essential in any country which makes claim to be recognized as a civilised community.

1.8.2.1.

Passage 28

NAME_____

DATE_____

WHAT IS LAW?

Read the following passage and complete it by filling in the blank spaces with the most suitable word you can think of. Only one word per blank space should be used. You have 15 minutes for this exercise.

Law has been defined as 'the just interference of the State in the interests and passions of humanity.' But the machinery of law has, in fact, changed (1)_____ all recognition in the past hundred years. It has (2)_____ to move with the times and has been modernized (3)_____ some extent, although it still has many imperfections. There (4)_____ also a genuine desire among the majority of lawyers (5)_____ assist in providing legal machinery which is more adequate (6)_____ the involved needs of modern life. Justice must, however, (7)_____ be uncertain no matter how lofty our intentions may (8)_____. Every Judge is himself John Citizen a member of (9)_____ community, subject to the same limitations of human intelligence, (10)_____ all the failings of our common lot.

It is (11)_____ the function of a Judge to decide a legal (12)_____ until after the facts have been presented and analyzed. (13)_____ enable you to form a judgement,

but as your (14)_____ reaction may be totally different from yours, you may (15)_____ arrive at a totally different judgement. If the Judge (16)_____ the evidence of your opponent, and you are the (17)_____ loser, you may leave the court murmuring that there (18)_____ no such thing as justice. This is very human (19)_____ very unreasonable. There is a loser to every lawsuit, (20)_____ although there are many dishonest people, the majority of (21)_____ are contested by parties each of whom believes himself (22)_____ be honest.

So it is possible to have the (23)_____ where two judges may honestly arrive at two different (24)_____ based on identical facts and no man can say (25)_____ of them, if either, is right. One must recognize that the most honest legal system can never be infallible.

1.8.2.2.

PASSAGE 2A : ACCEPTABLE WORD RESPONSES

Item No.	EXACT WORD	ACCEPTABLE WORDS
1	necessary	right, just, legitimate
2	the	a, its, our
3	would	could, might
4	or	and
5	majority	
6	amount	degree
7	the	
8	as	
9	building	structure, house
10	removed	dismantled
11	hope	call, opt, look, vote, wait, try, ask, prepare, plead, arrange
12	this	that, it
13	with	
14	cycling	
15	you	
16	liberty	freedom
17	later	
18	without	
19	because	
20	the	every, a
21	irresponsibility	maniacs, wrongdoers, individuals, eccentrics
22	because	as, since, for

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
23	as	
24	individual	
25	other	individual
26	rule	government, code, policeman, machinery, rules
27	however	unfortunately
28	is	becomes
29	offence	act, behaviour
30	any	a, this, that, the
31	made	legalised, formalised, enforced, formed, approved, established
32	as	
33	thirteenth	
34	Parliament	
35	one	some
36	the	
37	the	
38	read	presented, discussed, proposed, examined, considered
39	and	
40	receive	obtain, get
41	law	
42	may	can
43	the	
44	the	as
45	by	
46	represent	embody, be, express, reflect

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
47	applicable	applied
48	the	those
49	the	a, our
50	courts	court, judge, judges
51	the	
52	other	alternate, different
53	must	will
54	before	
55	given	delivered, made, passed
56	invariably	certainly, always, automatically,
57	the	
58	and	or
59	is	
60	and	which
61	established	new, introduced, instituted
62	trial	hearing, decision, judgement
63	reference	referring
64	sake	
65	judge	
66	called	
67	was	
68	involved	contained
69	deliver	make, pass, give
70	precedent	
71	can	
72	expression	term, English, so-called

ITEM NO	EXACT WORD	ACCEPTABLE WORDS
73	distinguish	describe, mean
74	from	
75	Statute	
76	says	
77	intends	wants, wishes
78	to	
79	everyone	he
80	for	
81	a	
82	are	
83	is	
84	with	
85	this	it
86	you	
87	going	
88	hearing	judgement, trial
89	as	
90	the	
91	if	
92	why	
93	no	without
94	and	but
95	law	
96	or	and
97	are	
98	shortcomings	imperfections, failings, weakness, faults
99	just	fair, perfect
100	is	

1.B.2.2. PASSAGE 2B : ACCEPTABLE WORD RESPONSES

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
1	beyond	
2	tried	tended, continued, had
3	to	
4	is	seems
5	to	
6	to	for
7	always	still, sometimes
8	be	
9	the	a, our, his
10	with	and, including
11	not	never
12	question	suit, dispute, case, issue, matter
13	facts	they
14	neighbour's	opponent's
15	each	both, ultimately, possibly, subsequently, sometimes
16	prefers	accepts, believes
17	disgruntled	unfortunate, eventual, unhappy, dissatisfied
18	is	
19	but	though
20	and	for
21	actions	cases, lawsuits
22	to	
23	situation	case
24	opinions	judgements, conclusions, decisions
25	which	

1.B.3.1.

Passage 3A

URBANISATION

(Text of Summary Experiment)

It is generally agreed that the first true cities appeared about 5000 years ago in the food-producing communities of the Middle East. The cities of Sumeria, Egypt and the Indus Valley possessed a number of characteristics which distinguished them as truly urban. The cities were very much larger and more densely populated than any previous settlement, and their function was clearly differentiated from that of the surrounding villages. In the cities the old patterns of kinship relations were replaced by a complex hierarchy of social classes based on the specialisation of labour. The cities acquired the basis for an effective capital by collecting the surplus of the primary producers in the villages. Moreover, the need to keep records led to the development of writing and arithmetic, and the increased sophistication of urban society gave a new impetus to artistic expression of every kind.

When the basis of city life was established in Europe the urban tradition was drawn from the ancient cities of the Middle East, via the civilizations of Greece and Rome. We can trace three main phases in the growth of the West European city. The first of these is the medieval phase, which extends from the beginning of the eleventh century to about 1500. The second is the Renaissance and Baroque phase which can be traced from about 1500 to the beginning of the nineteenth century. The third is the modern phase, extending from the early nineteenth century to the present day.

Every medieval city began as a small settlement which grew up round a geographical or cultural focal point. This would often be a permanent structure such as a stronghold, a cathedral or a large church. In districts where travel and trade were well established, it might be a market, a river crossing, or a place where two or more trade routes converged. In studies of urban geography the oldest part of a town is referred to as the nuclear settlement. There are many small towns in Europe where it is still possible to trace the outline of the original nuclear settlement. It is, of course, much more difficult to do this in the case of a large modern city which has grown to many times its original size.

From the point of view of the urban geographer, a city is essentially a group of dwellings arranged in such a way that the inhabitants can share in a variety of co-operative enterprises. All urban settlements must meet a number of basic requirements. They must be reasonably compact in form, so that all parts of the town are easily accessible. There must be adequate space between the buildings for pedestrian or vehicular traffic, and special areas or buildings must be set aside for public functions. Perhaps most important of all, the citizens must have security to enable them to carry on their affairs. For this reason the earliest settlements were often attached to an existing castle. At a later stage, when the urban settlements were larger and more presperous, the citizens usually surrounded the towns with walls of their own. Towns founded in the later Middle Ages were almost invariably surrounded by a strong wall. Each town had become, in effect, a fortress in itself.

The decision to establish a settlement in a particular place

depended basically on two factors: politico-cultural and economic. The builders of a fortress or important church were motivated primarily by politico-cultural considerations. These builders sought a prominent hill site or a promontory surrounded on three sides by sea, river or marsh. Such a site dominated the surrounding countryside and had the benefit of natural defences. On the other hand, an economic community required, above all, easy access, room to expand, and contact with the main trade routes. For this reason, a particularly favoured urban site was on fairly flat land by a navigable waterway. The two primary functions, politico-cultural and economic, were often combined in a single settlement which included both a hilltop and low-lying ground. Thus, in the old Greek cities, the Acropolis lay on the hilltop and the town on the lower slopes. In France and Belgium we find the same combination of the Haute-Ville and the Basse-Ville, which in Germany are called the Oberstadt and the Unterstadt. In each case the cathedral or secular stronghold, (or both), is situated on a hill, and the main part of the urban community is in the valley below.

There was a tendency in the Middle Ages for the same ideas to be repeated again and again. As a result we can distinguish not only certain recurring systems of urban planning, but also families of towns which have the same basic features of design.

Although some cities grew greatly in size during the Renaissance period, the metropolitan city as we know it today has its roots in the industrial revolution. Up to that time the process of urbanisation had affected only a small minority of the population. From the beginning of the nineteenth century, however, three major economic factors led to the growth of cities on an entirely new scale. Firstly,

the invention of powerful new machines gave rise to factories of unprecedented size, which created an enormous demand for labour. Secondly, the large-scale construction of roads, railways and canals provided cheap and regular transport which made possible the concentration of industries and population into particular areas. Thirdly, a revolution in agriculture led to the development of an efficient system of mixed farming, new methods of breeding and an increase in the yield of corn, all of which helped to provide the food necessary to sustain a greatly increased urban population.

The process of urban growth is still closely linked to industrial development, but the increased complexity of administration and commerce has also contributed to the rapid rate of urbanisation. It has been estimated that in 1800, less than three per cent of the world population, or 27.4 million, lived in towns of over five thousand inhabitants. By 1950 the proportion of town-dwellers had grown to nearly thirty per cent, or over 716 million. Moreover, during the last half-century, it is the larger cities that have succeeded in attracting the population. And with urbanisation has come the accusation that it helps breed a form of society that lacks values and standards of behaviour.

1.B.3.2

Passage 3A

NAME_____

DATE_____

URBANISATION

Read the following passage and complete it by filling in the blank spaces with the most suitable word you can think of. Only one word per blank space should be used. You have 45 minutes for this exercise.

It is generally agreed that the first true cities appeared about 5000 years ago in the food-producing communities of the Middle East. The cities of Sumeria, Egypt and the Indus Valley (1)_____ a number of characteristics which distinguished them as truly (2)_____. The cities were very much larger and more densely (3)_____ than any previous settlement, and their function was clearly (4)_____ from that of the surrounding villages. In the cities (5)_____ old patterns of kinship relations were replaced by a (6)_____ hierarchy of social classes based on the specialisation of (7)_____. The cities acquired the basis for an effective capital (8)_____ collecting the surplus of the primary producers in the (9)_____. Moreover, the need to keep records led to the (10)_____ of writing and arithmetic, and the increased sophistication of (11)_____ society gave a new impetus to artistic expression of (12)_____ kind.

When the basis of city life was established (13)_____ Europe, the urban tradition was drawn from the ancient

(14)_____ of the Middle East, via the civilizations of Greece (15)_____ Rome. We can trace three main phases in the (16)_____ of the West European city. The first of these (17)_____ the medieval phase, which extends from the beginning of (18)_____ eleventh century to about 1500. The second is the (19)_____ and Baroque phase which can be traced from about (20)_____ to the beginning of the nineteenth century. The third (21)_____ the modern phase, extending from the early nineteenth century (22)_____ the present day.

Every medieval city began as a (23)_____ settlement which grew up round a geographical or cultural (24)_____ point. This would often be a permanent structure such (25)_____ a stronghold, a cathedral or a large church. In (26)_____ where travel and trade were well established, it might (27)_____ a market, a river crossing, or a place where (28)_____ or more trade routes converged. In studies of urban (29)_____ the oldest part of a town is referred to (30)_____ the nuclear settlement. There are many small towns in (31)_____ where it is still possible to trace the outline (32)_____ the original nuclear settlement. It is, of course, much (33)_____ difficult to do this in the case of a (34)_____ city which has grown to many times its original (35)_____.

From the point of view of the urban geographer, (36)_____ city is essentially a group of dwellings arranged in (37)_____ a way that the inhabitants can share in a (38)_____ of co-operative enterprises. All urban settlements must meet a

(39)_____ of basic requirements. They must be reasonably compact in (40)_____ so that all parts of the town are easily (41)_____. There must be adequate space between the buildings for (42)_____ or vehicular traffic, and special areas or buildings must (43)_____ set aside for public functions. Perhaps most important of (44)_____, the citizens must have security to enable them to (45)_____ on their affairs. For this reason the earliest settlements (46)_____ often attached to an existing castle. At a later (47)_____, when urban settlements were larger and more prosperous, the (48)_____ usually surrounded the towns with walls of their own. (49)_____ founded in the later Middle Ages were almost invariably (50)_____ by a strong wall. Each town had become, in (51)_____, a fortress in itself.

The decision to establish a (52)_____ in a particular place depended basically on two factors: (53)_____ and economic. The builders of a fortress or important (54)_____ were motivated primarily by politico-cultural considerations. These builders sought (55)_____ prominent hill site or a promontory surrounded on three (56)_____ by sea, river or marsh. Such a site dominated (57)_____ surrounding countryside and had the benefit of natural defences. (58)_____ the other hand, an economic community required, above all, (59)_____ access, room to expand, and contact with the main (60)_____ routes. For this reason, a particularly favoured urban site (61)_____ on fairly flat land by a navigable waterway. The (62)_____ primary functions,

politico-cultural and economic, were often combined in (63)_____ single settlement which included both a hill-top and low-lying (64)_____. Thus, in the old Greek cities, the Acropolis lay (65)_____ the hill-top and the town on the lower slopes. (66)_____ France and Belgium we find the same combination of (67)_____ Haute-Ville and the Basse-Ville, which in Germany are called (68)_____ Oberstadt and the Unterstadt. In each case the cathedral (69)_____ secular stronghold, or both, is situated on a hill, (70)_____ the main part of the urban community is in (71)_____ valley below.

There was a tendency in the Middle (72)_____ for the same ideas to be repeated again and (73)_____ in the establishment of different towns. As a result (74)_____ can distinguish not only certain recurring systems of urban (75)_____, but also families of towns which have the same (76)_____ features of design.

Although some cities grew greatly in (77)_____ during the Renaissance period, the metropolitan city as we (78)_____ it today has its roots in the industrial revolution. (79)_____ to that time the process of urbanisation had affected (80)_____ a small minority of the population. From the beginning (81)_____ the nineteenth century, however, three major economic factors led (82)_____ the growth of cities on an entirely new scale. (83)_____, the invention of powerful new machines gave rise to (84)_____ of unprecedented size, which created an enormous demand for (85)_____. Secondly, the large-scale construction of

roads, railways and canals (86)_____ cheap and regular transport which made possible the concentration (87)_____ industries and population into particular areas. Thirdly, a revolution (88)_____ agriculture led to the development of an efficient system (89)_____ mixed farming, new methods of breeding and an increase (90)_____ the yield of corn, all of which helped to (91)_____ the food necessary to sustain a greatly increased urban (92)_____.

The process of urban growth is still closely linked (93)_____ industrial development, but the increased complexity of administration and (94)_____ has also contributed to the rapid rate of urbanisation. (95)_____ has been estimated that in 1800, less than three (96)_____ of the world population, or 27.4 million, lived in (97)_____ of over 5000 inhabitants. By 1950, the proportion of (98)_____ had grown to nearly thirty percent or over 716 (99)_____. Moreover, during the last half-century, it is the larger (100)_____ which have succeeded in attracting the population. And with urbanisation has come the accusation that it helps breed a form of society that lacks values and standards of behaviour.

1.B.3.2.

Passage 38

NAME_____

DATE_____

URBANISATION

Read the following passage and complete it by filling in the blank spaces with the most suitable word you can think of. Only one word per blank space should be used. You have 15 minutes for this exercise.

The typical modern metropolis has been described as a concentration of at least 500,000 people living within an area in which the travelling time from the outskirts to the centre is about 40 minutes. The metropolis has four major components: a central business (1)_____; a collection of manufacturing and allied industries; a quantity (2)_____ housing with its attendant services; and an area of (3)_____ land. The central business complex is made up of (4)_____ retail businesses, financial institutions and offices of the public (5)_____. A generation ago it was usual to find factories (6)_____ with business premises for space in the central area, (7)_____ the present tendency is for manufacturing industries to move (8)_____ the outskirts of the city where land is cheaper. (9)_____ accounts for the largest amount of occupied land in (10)_____ metropolis, and also presents the greatest problems in the (11)_____ of slums, or sub-standard dwellings, and the segregation of (12)_____

by income or race. The fourth major component of (13)_____ metropolis is open land maintained for recreational use which (14)_____ currently of great concern to urban planners.

It has (15)_____ claimed that with the development of the modern metropolis (16)_____ city has undergone a qualitative change. It is no (17)_____ merely a larger version of the traditional city, but (18)_____ entirely new form of settlement. Moreover, it is a (19)_____ which may contain within itself the seeds of its (20)_____ decay. Perhaps there is more substance in the argument (21)_____ the modern metropolis will eventually be choked to death (22)_____ its own growth. Certainly, an indispensible requirement of the (23)_____ will be careful planning with the aim of achieving (24)_____ rational distribution of the urban population, and of their (25)_____ activities. Another requirement will be the development of cheap and efficient public transport facilities to connect the various parts of the metropolis.

1.B.3.3.

PASSAGE 3A : ACCEPTABLE WORD RESPONSES

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
1	possessed	shared, had
2	urban	civilised, different, urbanised
3	populated	
4	differentiated	different, separate, distinct
5	the	
6	complex	new, distinct
7	labour	work, function, functions
8	by	
9	villages	neighbourhood, area, countryside
10	development	invention, practice, study
11	urban	this
12	every	some
13	in	
14	cities	communities, settlements
15	and	
16	growth	development, rise, history
17	is	
18	the	
19	Renaissance	Post-Medieval
20	1500	
21	is	
22	to	until
23	small	nuclear, embrionic
24	focal	strategic, central

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
25	as	
26	districts	places, areas, cities, settlements, towns
27	be	
28	two	
29	geography	development, growth, society
30	as	
31	Europe	
32	of	
33	more	
34	modern	medieval, old, new
35	size	
36	a	the
37	such	
38	variety	number, lot
39	number	minimum
40	form	size, structure, themselves
41	accessible	reached, connected
42	pedestrian	people
43	be	
44	all	
45	carry	
46	were	are
47	stage	date, period, time
48	citizens	inhabitants, people
49	towns	cities

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
50	surrounded	
51	effect	fact, essence
52	settlement	city, town
53	politico-cultural	
54	church	castle, stronghold
55	a	
56	sides	
57	the	
58	on	
59	easy	free
60	trade	trading, commercial, traffic
61	was	
62	two	
63	a	one
64	ground	land
65	on	
66	in	
67	the	
68	the	
69	or	
70	and	while, whereas
71	the	
72	Ages	
73	again	
74	we	one, you
75	planning	growth, development, design

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
76	basic	outward, principal, main, particular
77	size	
78	know	see, recognise
79	up	
80	only	
81	of	
82	to	
83	firstly	
84	factories	industries
85	labour	space, workers, food, land
86	provided	offered, permitted, created
87	of	
88	in	
89	of	
90	in	
91	provide	produce, supply
92	population	
93	to	with
94	commerce	industrialisation, economics
95	it	
96	percent	
97	towns	cities
98	town-dwellers	inhabitants
99	million	
100	cities	town

1.B.3.3.

PASSAGE 38 : ACCEPTABLE WORD RESPONSES

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
1	complex	area, sector
2	of	
3	open	park
4	diversified	various, many, large, some, several, the
5	administration	sector
6	competing	fighting, vying
7	but	now, however, whereas
8	to	toward, towards
9	housing	
10	the	
11	form	increase, growth
12	people	
13	the	any, a
14	is	
15	been	
16	the	
17	longer	more
18	an	
19	form	settlement, growth, pattern
20	own	
21	that	
22	by	
23	future	
24	a	
25	various	different, many, specific, differing, varying

1.B.4.1.

Passage 4A

NAME_____

DATE_____

THE LEATHERBACK TURTLE

Read the following passage and complete it by filling in the blank spaces with the most suitable word you can think of. Only one word per blank space should be used. You have 45 minutes for this exercise.

A leatherback turtle emerging from a phosphorescent surf on a tropical beach is an awesome creature. It is by far the largest turtle in the (1)_____, outweighing the most ponderous of the giant Galapagos tortoises (2)_____ several hundred pounds. The average weight of the leatherbacks (3)_____ in French Guiana, where we tried to weigh them, (4)_____ about 800 pounds; the biggest probably weigh up to (5)_____ pounds.

No accurate surveys have been made because hoisting (6)_____ creatures on scales in the field is a monumental (7)_____. We brought with us a spring-pulley scale and tried (8)_____ haul a couple of average-looking leatherbacks up on a (9)_____ system made from freshly cut saplings. But each time (10)_____ rope broke before the turtle was lifted even partly (11)_____ the sand.

We had come to French Guiana under (12)_____ auspices of the World Wildlife Fund to tag, measure (13)_____ study as many leatherbacks as possible during a few (14)_____ of

the summer of '74. It wasn't until 1968 (15)_____ Pritchard discovered this remarkable stretch of beach as a (16)_____ nesting site for these turtles. Starting in 1969, he (17)_____ tagging groups which camped on the beach each summer (18)_____ several weeks to mark and record turtles. This beach (19)_____ now considered the single most important known nesting site (20)_____ the leatherback.

The leatherback, as its name implies, is (21)_____ distinguished by its shell. Instead of having a hard (22)_____ composed of separate plates like other sea turtles, the (23)_____ has a leathery exterior with seven bony ridges running (24)_____ length of its back, and five along its belly. (25)_____ in its skin is a mosaic of small, flat (26)_____ forming a hard, inner casing that, together with a (27)_____ of oily cartilage sometimes as much as two inches (28)_____, helps do the job of a shell. The leatherback's (29)_____ is also quite distinctive: flippers, head and back are (30)_____ grey-black, irregularly spotted with small, whitish blotches.

There are (31)_____ reasons why the leatherback is considered unique. Fragmentary evidence (32)_____ that it wanders farther in the oceans and can (33)_____ greater variations in water temperature than other sea turtles. (34)_____ jaw construction is amazingly fragile so that hard-shelled organisms (35)_____ other turtles crush easily, molluscs and crustaceans, could injure (36)_____ leatherback. Lacking the massive construction and crushing plates of (37)_____ other sea turtles, leatherbacks seem to rely

almost exclusively (38)_____ a diet of jellyfish.

The leatherback's skeleton is oddly (39)_____. It has large amounts of cartilage at the ends (40)_____ the bones, thereby theoretically possessing a potential for unlimited (41)_____. Its skull is also a complex conglomeration of small (42)_____ which completely disassemble after the flesh is removed. Numerous (43)_____ characteristics separate the leatherback from its better-known cousins so (44)_____ some herpetologists believe that the leatherback deserves a sub-order (45)_____ its own and refer to it 'as the fly (46)_____ the ointment of turtle classification.'

The leatherback is also (47)_____ in that it does not live long in captivity. (48)_____ Ceylon, P.E.P. Deraniyagala was able to keep a hatchling alive (49)_____ two years back in the 1930's while at the (50)_____ Seaquarium a young leatherback had reached thirteen months by (51)_____ August. Leatherbacks usually bang their noses against the sides (52)_____ any tank in which they are put until they (53)_____ of head injuries. Even in death, the leatherback doesn't (54)_____ into a system. Because its whole body is suffused (55)_____ oil which is difficult to remove, even 'dry' preserved (56)_____ in museums drip for years so that they eventually (57)_____ to be discarded.

Leatherbacks encounter as many hazards on (58)_____ natural forays ashore as they do in aquariums. On (59)_____ tagging journeys through the area of lifeless trees, swamp (60)_____ mangrove barrier, which stretched along a section we called (61)_____ 'dead forest', we would rarely

find nesting leatherbacks. Turtles (62)_____ usually unable to penetrate the driftwood and have to (63)_____ along to where the driftwood breaks up into a (64)_____ nesting area. At two or three places, however, there (65)_____ a passage just wide enough to allow a turtle (66)_____ crawl through.

But turtles don't follow their own tracks (67)_____ as their guidance systems on land seem to be (68)_____ to locating the general direction of the sea. So (69)_____ a leatherback unlucky enough to penetrate the barrier finishes (70)_____, she turns to the sea only to run again (71)_____ the driftwood. This time she probably will not make (72)_____ through. On several occasions we found leatherbacks dead from (73)_____ and exposure after having pulled their enormous weights hundreds (74)_____ yards along the beach in a desparate search for (75)_____ opening. These unfortunate creatures were close enough to have (76)_____ the spray of the surf, yet were unable to (77)_____ the sea.

Sometimes we would find live turtles vainly (78)_____ about on the land side of the barrier. Many (79)_____ these we were able to guide back to the (80)_____ by waving a flashlight from the direction we wished (81)_____ to take since leatherbacks seem to follow a light (82)_____ instinctively. But some turtles were beyond our help, either (83)_____ they were hopelessly entangled or because a suitable opening (84)_____ the barrier was too far away for them in (85)_____ weakened condition. A leatherback

cannot reverse its direction; it (86)_____ committed to forward movement, though some turning is possible. (87)_____ if a leatherback runs into a solid barrier, it (88)_____ not be able to extricate itself.

During the nights (89)_____ the turtles came ashore to nest, we would tag (90)_____ from ten to eighty of them. In peak months (91)_____ previous seasons as many as hundred and forty were (92)_____ in one night. A leatherback tagged while actually laying (93)_____ reacts very little. The tagging does no damage to (94)_____ turtle for even if the small wound does bleed, (95)_____ leatherback possesses the peculiar ability to shut off the (96)_____ of blood to an extremity until healing occurs. The (97)_____ apparently nests in alternate years, though the validity of (98)_____ absolute statement is far from certain. Where they come (99)_____ or where they go after nesting no one knows (100)_____ sure although one can assume that the leatherback travels (101)_____. The first tag return came from a turtle tagged (102)_____ 1970 across the Maroni River in Surinam near French (103)_____ that turned up off the coast of Ghana, West (104)_____, almost a year later.

All sea turtles use navigational (105)_____ that are still not understood by scientists. Perhaps some (106)_____ our tag returns will aid in understanding how these (107)_____ can travel 3000 miles or more in the open (108)_____ and yet return to nest on the same small (109)_____ time and time again.

Their life in the water (110)_____ a mystery. While other sea turtles are found regularly (111)_____ offshore feeding grounds, the leatherback just disappears. Many things (112)_____ us to believe that this turtle lives in the (113)_____ ocean. One thing is sure, the cycle of the (114)_____ sea turtle has worked successfully since long before the (115)_____ species was evolved. So far the leatherback has not suffered serious damage from humans, but the increase in oceanic pollution and the encroachment of civilisation on the nesting grounds may yet drive this strange giant into extinction.

1.B.4.1.

Passage 4B

NAME_____

DATE_____

THE LEATHERBACK TURTLE

Read the following passage and complete it by filling in the blank spaces with the most suitable word you can think of. Only one word per blank space should be used. You have 15 minutes for this exercise.

We came to French Guiana under the auspices of the World Wildlife Fund to tag, measure and study as many leatherback turtles as possible during a few weeks of the summer of '74. Though they roam the world's oceans, their nesting concentrations (1)_____ few and cover only small geographic areas. The leatherback's (2)_____ precarious situation has spurred efforts to study and protect (3)_____ Archie Carr, the University of Florida turtle authority, has (4)_____ 'this extraordinary turtle, one of the most remarkable of (5)_____ living reptiles.' The only detailed studies made of the (6)_____'s habits so far have been done when the female (7)_____

The female prepares the beach for nesting by first (8)_____ great sweeps with her front flippers, excavating a pit (9)_____ to three yards wide, according to the span of (10)_____ flippers. When finished, the nest will have a bell-shaped (11)_____ two to three feet deep into which the turtle (12)_____ drop several dozen soft-shelled eggs, the average clutch in (13)_____ Guiana usually numbering around

eighty-six eggs. Another peculiarity of (14)_____ leatherback is that its normal eggs are always accompanied (15)_____ a number of small yolkless eggs which have not (16)_____ been adequately explained. And after the laying is finished, (17)_____ will cover the egg cavity with sand pushed in (18)_____ her hind flippers.

The whole nesting process takes about (19)_____ hours, after which the turtle heads wearily for the (20)_____. When the waves touch her, the leatherback's energy is (21)_____. Leatherbacks, so heavy and clumsy on land, are perfectly (22)_____ to the marine environment. Flippers that flopped about on (23)_____ move in the water with a smooth style that (24)_____ the flight of a hawk. Its flippers are proportionately (25)_____ enabling it to make faster swims over longer distances. Carr calls the leatherback 'the most aquatic of all turtles and, indeed, the most completely pelagic of modern reptiles except sea snakes or, perhaps, the Loch Ness Monster.'

1.8.4.2.

PASSAGE 4A : ACCEPTABLE WORD RESPONSES

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
1	world	
2	by	of
3	nesting	living, found, there, down
4	is	was
5	1300	
6	these	those, the
7	task	job, problem
8	to	
9	tripod	pulley, weighing, lever, balance
10	the	
11	off	from
12	the	
13	and	
14	weeks	
15	that	
16	major	favourite, important, regular, ideal
17	led	organized
18	for	
19	is	was
20	for	of
21	immediately	easily, soon, always, characteristically
22	carapace	back, shell
23	leatherback	
24	the	

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
25	imbedded	also, there, down
26	bones	ridges, plates
27	layer	lot, sort, mass
28	thick	deep
29	coloration	colour, appearance, body
30	a	all, nearly, almost, usually
31	many	several, various, numerous
32	suggests	shows, reveals, indicates
33	tolerate	resist, stand, withstand, bear, survive
34	its	the, their
35	that	which
36	a	the
37	the	some, many, most
38	on	
39	immature	made, formed, built, constructed, composed
40	of	
41	growth	length
42	bones	
43	additional	other
44	that	
45	of	
46	in	
47	alone	unique, different
48	in	
49	for	over, almost

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
50	Miami	same, Floride
51	last	late, early, this, mid
52	of	
53	die	collapse
54	fit	
55	with	
56	specimens	ones, leatherbacks, turtles
57	have	
58	their	
59	our	the, these
60	and	
61	the	
62	are	
63	pass	swim, go, travel, crawl, move
64	better	suitable, easier, good
65	was	is
66	to	
67	back	easily, often, again
68	limited	confined, restricted
69	when	once
70	nesting	
71	into	against
72	it	
73	exhaustion	tiredness, hunger
74	of	
75	another	an, some

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
76	felt	smelt
77	reach	enter
78	flopping	crawling, moving
79	of	
80	sea	water
81	them	
82	source	flash, almost, quite
83	because	
84	in	
85	their	a, this
86	is	
87	thus	therefore, so, even
88	may	will, might
89	when	that, as
90	anywhere	sometimes
91	of	in, during
92	tagged	marked, seen, labelled
93	eggs	
94	the	a
95	the	a
96	flow	supply, circulation
97	leatherback	turtle
98	any	an
99	from	
100	for	
101	far	around, widely, fast

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
102	in	
103	Guiana	
104	Africa	
105	techniques	methods, systems
106	of	
107	creatures	leatherbacks, turtles, reptiles
108	ocean	sea
109	beach	place, island, site
110	remains	is
111	in	near, around, at
112	lead	force
113	open	deep, great
114	leatherback	deep, great
115	human	

1.8.4.2.

PASSAGE 4B : ACCEPTABLE WORD RESPONSES

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
1	are	
2	potentially	present, recent, very, current
3	what	
4	called	
5	all	our
6	leatherback's	turtle's
7	nests	
8	making	digging
9	two	up
10	her	the, its
11	cavity	hole, hollow, pit
12	will	
13	French	
14	the	this
15	by	
16	yet	
17	she	it
18	by	with
19	two	three
20	water	sea, ocean
21	renewed	regained, restored, revived
22	adapted	adjusted, accommodated, accustomed, used, suited

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
23	shore	land, sand
24	recalls	resembles, emulates, equals
25	larger	built, made, big, strong, sized, formed, large

1.B.5.1.

Passage 5A

NAME_____

DATE_____

SCIENTIFIC STUDY OF SOCIAL DEVELOPMENT

Read the following passage and complete it by filling in the blank spaces with the most suitable word you can think of. Only one word per blank space should be used. You have 45 minutes for this exercise.

History, as it is commonly understood, relates to the behaviour of groups of human beings, identified by certain qualities that remain relatively unchanged over an extended period of time. A scientist, however, experiments in the laboratory with a (1)_____ of matter under conditions he can control; and the (2)_____ so derived enables him to analyze the history of (3)_____ physical processes as they unroll themselves uncontrolled by him. (4)_____ can thus make certain restricted historical predictions, such as (5)_____ of eclipses, with great certainty.

On the other hand (6)_____ also changes the course of history, usually unconsciously, by (7)_____ design and creation of technical processes that profoundly affect (8)_____ life. Now, while he can predict exactly how much (9)_____ produce a machine designed on his principles will discharge, (10)_____ how efficient it will be in production, he cannot (11)_____ what will be the social outcome of the machine (12)_____ has designed. Neither

can he tell how it will (13)_____ the standard of life of the community or what (14)_____ or intellectual movements it may stimulate. Did the inventors (15)_____ steam power forsee the devastation that would be wrought (16)_____ the mining areas of South Wales? Or did the (17)_____ of moving pictures anticipate the details of a new (18)_____ art? That type of social prediction is completely outside (19)_____ scope as an experimental scientist.

Actually, the operation of (20)_____ machine he has designed, forcing a new consumption of (21)_____ material and an altered distribution of finished products, is (22)_____ an experiment that changes the lives of men and (23)_____, but it is not one that he is concerned (24)_____ or even tries to control. The material operated upon, (25)_____ human beings in their social setting, has not been (26)_____ subject of his study. He has no technique for (27)_____ upon it. And yet, in a restricted sense, he (28)_____ make history, control its making, and predict its happening. (29)_____ does this within the four walls of his laboratory.

(30)_____ history of the past and of the present is (31)_____ made under conditions that approximate to those in a (32)_____ laboratory. True, both human beings and inanimate matter make (33)_____ histories in the material environment handed down to them (34)_____ the past. In one case it is 'human nature' (35)_____ man, and in the other, 'properties' for objects. But (36)_____ it is possible to detect the pattern in the (37)_____ of

inanimate matter by deliberately 'making' and 'repeating' experiments (38)_____ the laboratory, in the case of human history it (39)_____, in the first place, by an analysis of actual (40)_____ uncontrolled events that there can be any hope of (41)_____ a pattern.

There is yet another difference. For most (42)_____, sciences like astronomy and geology could be investigated because (43)_____ the recurrent nature of things observed, e.g. the motion of (44)_____ bodies. The story of man is different because each (45)_____ he makes in his history alters his environment, each (46)_____ in his environment changes that in which his fellow-men (47)_____ the oncoming generation live. The background of the parent (48)_____ from that of the child so that in a (49)_____ social atmosphere, men themselves react differently, acquiring new knowledge, (50)_____ experience, new understanding. They value things anew and human (51)_____ is itself transformed. So the clock of history never (52)_____ twice the same, it is not a mechanical process.

(53)_____ such a situation of continual flux where, then, are (54)_____ defining qualities that must help us if we wish (55)_____ understand this developing situation? For this purpose let us (56)_____ the characteristics of man, and of the material factors (57)_____ his environment just as is done in other fields (58)_____ science.

We note that in the first place, in (59)_____ to the food supply, which is their first essential, (60)_____

are gatherers, seeking and hunting their food, and consuming (61)_____ immediately. As against this, human beings are primarily food (62)_____. Planning their activities in advance towards this end, they (63)_____ their accumulated store of knowledge, the experience of their (64)_____, to create tools and devise schemes in order the (65)_____ to produce their food.

There follows a very significant (66)_____. Where animals are gregarious, living in herds, they are (67)_____ from dispersing mainly because of the need for protection (68)_____ outside attack and in the sequel, build up internally (69)_____ complicated species of society. Where food abounds to satisfy (70)_____ wants, the next strongest animal need then becomes the (71)_____ that binds the society together. This has been clearly (72)_____ and explored by Zuckerman in his Social Life of (73)_____. There, sex mates are conserved and the wives of (74)_____ strongest monkey or ape become his private property.

The (75)_____ of such a society, however, is profoundly different from (76)_____ of a food producing community of human beings. The (77)_____ fact of production rather than collection and direct consumption (78)_____ an active, planned attack on the soil, the changing (79)_____ the environment in the neighbourhood of the community; and, (80)_____ the group has surpassed the merely nomadic and pastoral (81)_____, the establishment of settlements leading to farmsteads and villages. (82)_____'s struggle with nature is an effort to use and

(83)_____ it. His objectives are different from those of animals. (84)_____ the course of his social evolution shows a fundamentally (85)_____ trend. It would therefore be ridiculous to see any (86)_____ parallel between the laws of animal communities and those (87)_____ human societies for while the animal seeks only immediate (88)_____, man schemes for security in the future, for later (89)_____, renouncing the present for the sake of the future. (90)_____ he begins to esteem <u>future</u> as against <u>present</u> use. (91)_____ it is important to safeguard the future, it is (92)_____ and desirable to sacrifice some of the present to (93)_____ end. He forgoes the 'lower' for the sake of (94)_____ 'higher'. He evolves a social code of an ethical (95)_____. He strives not simply towards survival but towards development.

(96)_____ the story of the struggle for social betterment this (97)_____ is continually obscured. To justify the ferocity of economic (98)_____ in human society, for example, on the grounds that (99)_____ is no more than a universal law of nature, (100)_____ to offer a specious excuse for an anti-social bias. Thus while animals may wage a struggle for mere existence, human societies wage something on a different level, and the tools and technical methods he devises are an essential feature of this process.

1.B.5.1.

Passage 5B

NAME_____

DATE_____

SCIENTIFIC STUDY OF SOCIAL DEVELOPMENT

Read the following passage and complete it by filling in the blank spaces with the most suitable word you can think of. Only one word per blank space should be used. You have 15 minutes for this exercise.

The technical level, resources, and skill at the command of the community and the way in which they are used, provide the basis out of which developments emerge. While it is true to say that at any (1)_____, the tools and machinery and technical skill of man (2)_____ for the working of the natural resources at his (3)_____, it is true only in a static sense. In (4)_____ the process is a dynamic one. The interaction of (5)_____ tools with the natural resources results in continual change (6)_____ one and adaptation in the other. As the tools (7)_____ natural resources, they in their turn provide for the (8)_____ of more complex machines and more elaborate tools, in (9)_____ rising levels, in ever increasing efficiency and in ever (10)_____ variety. In the wake of the discovery of coal (11)_____ iron, for instance, came steam power machines, transport and (12)_____ modes of communication, all in reality, elaborate tools.

Extended (13)_____ throws open more and more varied minerals from other (14)_____ of the earth. In a sense,

the boundaries of (15)_____ area occupied by man are continually extended. Science with (16)_____ geological theories, metallurgical and chemical inveetigations, is called into (17)_____. The fund of knowledge increases by leaps and bounds, (18)_____ with it, potential control over nature. The rise in (19)_____ level forces up in its turn the possible level (20)_____ social life, demanding of the community that it adapt (21)_____ to newer and newer levels of life and of (22)_____, while yet being itself forced into growth by the (23)_____ reaction it engenders, the expanding experience and needs of (24)_____ very members of the community itself. Science and technology (25)_____ from one side, the effects of social life. However on the other side, the potential level of social life is perpetually driven upwards by the very science and technology it produces; and the drive of technique is ultimately the main causal agency which, as it rises, accentuates and enriches also the quality of communal life.

1.B.5.2.

PASSAGE 5A : ACCEPTABLE WORD RESPONSES

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
1	portion	sample, specimen
2	information	result
3	actual	similar, some
4	he	
5	those	dates, times, occurrences
6	he	man
7	the	his
8	social	his, human
9	finished	actual
10	and	or
11	predict	know
12	he	
13	affect	improve, change, raise, influence
14	artistic	economic, social, political, cultural, ideas
15	of	
16	in	upon, amidst
17	inventors	inventor, production
18	screen	visual, plastic, creative, cinematographic
19	his	
20	the	each
21	raw	fresh, new
22	itself	also, really, equally, fundamentally
23	women	society

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
24	with	about
25	these	
26	the	a
27	experimenting	working, operating
28	does	can, helps, may, will
29	he	man
30	the	human
31	not	only
32	scientific	scientist's, experimental, science, small, research
33	their	
34	from	by
35	for	or
36	whereas	while, if, although, though
37	history	study, nature, case
38	in	within
39	is	starts
40	past	and, specific
41	seeing	discovering, detecting, finding
42	purposes	
43	of	
44	heavenly	celestial
45	step	decision, move
46	alteration	change, event
47	and	
48	differs	

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
49	changed	changing, different, new, given
50	new	fresh
51	nature	experience, society, life, history
52	strikes	rings
53	in	
54	the	these
55	to	
56	examine	study, define, consider, analyse
57	in	within, changing, constituting
58	of	
59	relation	regard, respect, connection
60	animals	they
61	it	
62	producers	growers, providers, cultivators
63	utilise	use, pooled, exploit
64	predecessors	ancestors, parents, past
65	better	
66	fact	difference, distinction, point
67	prevented	kept
68	against	from
69	a	
70	their	essential
71	factor	link, tie, one, fear, element, force
72	analysed	demonstrated, expressed, explained, examined, indicated, described, investigated

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
73	monkeys	apes, primates
74	the	
75	structure	organization, pattern, history, working
76	that	one
77	very	essential, mere, single, principal
78	argues	necessitates, means, indicates
79	of	
80	once	when, after
81	stage	phase
82	man's	
83	control	direct, subdue, master, change, exploit
84	accordingly	so, thus, and, therefore
85	different	
86	close	real, straightforward, rigorous, clear
87	of	
88	satisfaction	security, advantage
89	enjoyment	comfort, benefit, generations, use
90	hence	thus, so, accordingly, consequently
91	if	since, as, because
92	valuable	necessary, essential, advisable, reasonable
93	that	this
94	the	
95	nature	
96	in	

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
97	point	fact, striving, tendency
98	competition	struggle
99	it	competitiveness
100	is	

1.B.5.2.

PASSAGE 5B : ACCEPTABLE WORD RESPONSES

ITEM NO.	EXACT WORD	ACCEPTABLE WORDS
1	stage	period
2	suffice	
3	command	disposal
4	reality	fact
5	the	man's, those, his
6	in	
7	transform	use
8	possibility	development, emergence, invention, use
9	ever	continually
10	new	increasing, wider, greater, growing
11	and	
12	other	new, various
13	transport	exploration, investigation, technology
14	regions	parts, quarters
15	the	
16	its	
17	being	play, action
18	and	
19	technical	
20	of	
21	itself	continually
22	production	technology, science
23	human	chain, continual, very
24	the	
25	are	

1.C.

INFORMAL EXPERIMENT : NO-CONTEXT CLOZE VERSUS CONTEXT CLOZE

[In the experiment, each frame was on a separate sheet of paper.]

1. this stage, however, no major __technical__ advance, apart from the development ...
2. turn gave rise to more __complex__ and more skilled operations. Among ...
3. the weaver, potter and the __flint__ polisher. There was at this ...
4. 1000 years later. But wherever __it__ was adopted, agriculture had the ...
5. the most profound effects on __human__ life. Its primary function was ...
6. food, but in doing so __it__ made possible a vast number ...
7. than that. From the beginning __it__ was also a basis for ...
8. order and enforcing the law. __Cities__ were the centres, therefore, not ...
9. ruling class. The history of __ancient__ civilisations illustrates very clearly the ...
10. Indus in northern India, and __perhaps__ simultaneously the Yangtse-kiang in China. ...
11. no leap year, and this __calendar__ was probably first used in ...
12. of a good supply of __meat__ to use their leisure to ...

13. class. With the growth of ___cities___, various technical problems arose which ...

14. in 4242 B.C. Knowledge of ___astronomy___ and the working of the ...

15. food surplus is to support ___workers___ in crafts and industries other ...

16. an average of about 1350 ___kg___ of wheat was produced per ...

17. or in the form of ___finished___ products such as axes, vessels ...

18. cultivation. Cattle, sheep, goats and ___pigs___ were used earlier while the ...

19. The immediate consequence of these ___advances___ is illustrated in the Faiyum ...

20. the way in which the ___development___ of science is related to ...

I.D.

INFORMAL EXPERIMENT : ABBREVIATED TEXT

This is an abbreviated text. (1) Reconstruct the original text indicating (2) where, according to you, the missing item(s) were initially included.

The domestication of animals for food began at about same time as cereal cultivation. Cattle, sheep, goats and were used earlier while the use of the horse, ox and camel for transport and traction came later. Early stages of settled agriculture, with both cultivated cereals domestic herds, have been most studied at the ancient of the Faiyum and Merimde in Egypt, and Sumer Mesopotamia. It is at the latter that the use sheep's wool for clothing is first recorded, probably 4000 before the present era. Dairy farming also seems to begun in Mesopotamia, perhaps 1000 years later.

With the growth of various technical problems arose which had to be solved. Of them was the need to supply clean water ensure the disposal of waste. At the city of Mohenjo-daro, on the River Indus, there was a well-organized system aqueducts and drains, and there is evidence of town-planning the city government. Civilisation not only creates material wealth an elaborate technology, but by making leisure possible it to the development of the arts. The artists of primitive times took advantage of a good supply of to use their leisure to paint and draw on walls and on bone.

Are you a native speaker of English? Yes No

APPENDIX TWO

READABILITY FORMULAE

2.A. THE FOG INDEX (1952) : Robert Gunning

$$\left((sl + \left(\frac{\text{Words of 3 + syllables}}{\text{Total number of words}} \times 100 \right) \right) \times .4 = \text{Fog Index.}$$

sl = sentence length i.e. total number of words divided by total number of sentences. Do not count proper names, compounds and verb forms, that with -ed or -es become 3 syllables for the 3+ syllable count.

This index provides grade placement according to the American system of education. An index of 13 to 16 is equivalent to Undergraduate level, while 17 + is at Postgraduate level. This formula normally over-estimates.

2.B. SMOG (1969) Simple Measure of Gobbledygook : McLaughlin

(i) Select 30 sentences, ten each from the beginning, middle and end of the passage.

(ii) Count the words with more than 3 syllables in the sample of 30 sentences. This gives N.

(iii) Apply formula:

N + 3 = Grade Placement with 100% comprehension.

This formula normally under-estimates.

2.C. NEW READING EASE INDEX (1951) : Farr-Jenkins-Paterson

(i) Use either total sample or 100-word samples.

(ii) Count the number of sentences to give S, and the number of words to give W.

(iii) Count number of one-syllable words. This refers to Ease and gives NOSW.

(iv) Apply formula:

$$\left((1.5999 \times (\frac{NOSW}{W} \times 100) \right) - (1.015 \times \frac{W}{S}) - 31.517$$

= New Reading Ease Index.

2.D. MTUL (1966) Minimal Terminable Unit Length : Kellogg Hunt

(i) Count total number of words in sample.

(ii) Count number of main clauses in sample.

(iii) Apply formula:

Number of words ÷ number of main clauses = MTUL.

2.E. THE READABILITY GRAPH (1968) : Edward B. Fry

(i) Take three samples of 100 words each from the beginning, middle and end of the sample.

(ii) Count the total number of sentences, divide by 3 = average number of sentences in one sample.

(iii) Count total number of syllables, divide by 3 = average number of syllables in one sample.

(iv) Apply formula which is done by a graph. This has the number of sentences on the axis, and the number of syllables on the abscissa, and gives the grade placement, once again according to the American system.

2.F. CLOZE FORMULA (1974) : John Bormuth

(i) Take one 100 word sample. Count all words and

letters in the sample. This gives (LET ÷ W), i.e. number of letters per word.

(ii) Count the number of punctuation units in the sample. This gives (LET ÷ MPU), i.e. number of letters per minimal punctuation unit.

(iii) Apply formula:

$$1.051674 - .099691\ (\text{LET} \div \text{W}) - .004236\ (\text{LET} \div \text{MPU}) + .000015\ (\text{LET} \div \text{MPU})^2 = \text{Cloze Mean Score}$$

APPENDIX THREE

ORGANIZATION OF TEXT

3.A.

RHETORICAL ANALYSIS OF TEXT IN TERMS OF DISCOURSE FUNCTION

A rough rhetorical analysis of the two texts used in Experiment II is presented here. The primary purpose was to make explicit what we meant by structure and organization of text, and explain the procedure adopted for preparing the list of ideas used in the summary. Though the method adopted is crude, it does help to define the kind of operations we expect the non-native reader to perform as he proceeds fairly rapidly through the text if he is to appreciate the meaning of linguistic units in the context of use.

The overall approach owes much to an undated paper by L. Thomas[1] as do some of the terms. But for the sake of clarity, we have defined all the terms that have been used in our analysis in terms of their discourse function. In this connection, the word 'theme' has been used to refer to the central idea in the text as a whole. The word 'topic' has been used to describe the main idea of a smaller unit such as a paragraph or group of sentences. Hence in any self-contained text of any length there is likely to be one or perhaps two themes, while the number of topics is likely to be far greater.

There are eight terms in all used for the analysis.

assertion : this refers to a statement or proposition that occurs for the first time, and which is likely to be the theme or topic.

evidence : this is taken to refer to evidence in support of the proposition, and will of necessity accompany it.

example : this is taken to refer to illustration or example of the proposition or evidence, and will generally accompany them. By definition this will refer to specific detail rather than generalized comment.

summing up : this refers to summing up or result based on the general assertion, evidence and example, and will usually follow them, though it can on occasion serve to introduce the argument. This often gives the summary of the argument.

qualification: this refers to an explicit attempt by the author to modify an assertion or proposition he has made earlier. This qualification may be either a modification or may limit the application of the proposition in some way.

repetition : this category is reserved for the proposition that is reiterated, not just within the paragraph structure but more often at a later stage when the author wishes to establish or re-establish a link with an earlier assertion. The assignation of text to this category often calls for inferential recognition of redundancy, or appreciation of the value of the proposition being reinforced by use of this device.

elaboration : this term refers to information which expands the proposition or evidence. This usually comprises an additive aspect, often a paraphrase spelling out the assertion in greater detail.

link : this last category is really a rag-bag where the unit (or sentence) appears to have been inserted more for purposes of cohesion of text than for furthering the argument. Sentences which are classified as link will usually have some reference to old information in order to tie them in with the new information that follows. The main difference between repetition and link is in terms of function, the latter performing a bridging operation between different parts of the text and occurring at paragraph boundaries.

The analysis of the texts that follow proceed linearly, taking the sentence as the sub-unit of the paragraph. Where the sentence performs more than one rhetorical function, each is identified. Against these is the content of the unit analysed, and against each paragraph is the summary. It will be seen how in some cases the latter is abstracted from explicit statements in the text, and in others, is the result of inference.

As we said earlier this analysis provides the rationale for arriving at the list of ideas that were

used for judging the summary question in Experiment II.[2] It will be seen that what is labelled assertion and evidence provided most if not all the ideas.

NOTES: 1. L.F. Thomas: A Flow Diagram Technique for describing the "Meaning Structure" of a text (undated mimeograph) 11 pp.

2. The list of ideas is provided separately in Appendix 3.8.

3.A.1.

ANALYSIS OF TEXT : PASSAGE 1

Sentence No.	Rhetorical Analysis	Discourse Function of Content
PARAGRAPH I		
1.	assertion	main theme : earliest civilisations depended on existence of wild cereals
	example	Middle East

SUMMARY: Earliest civilisations grew up where wild cereals could be found.

PARAGRAPH II		
2.	assertion	link with paragraph 1 domestication of animals for food began along with cereal cultivation.
3.	example	cattle, sheep etc.
	assertion	animals used for transport and traction a little later.
	example	horse, ox, etc.
4.	repetition	(topic) settled agriculture
	summing up	(cultivated cereals and domestic herds) developed.
	example	Egypt etc.
5.	evidence	use of wool for clothing.
6.	evidence	dairy farming began.

Sentence No.	Rhetorical Analysis	Discourse Function of Content

SUMMARY: Domestication of animals for food.
Animals used for transport and traction.
Use of wool for clothing, and dairy farming begun.

PARAGRAPH III		
7.	assertion	topic : agriculture had profound impact on life
	link	(logical result) (link with paragraph 2 which contained the evidence)
8.	evidence	purpose of agriculture to grow more food to enable new activities that changed man from a helpless to a dominating creature.
9.	evidence	size of agricultural communities increased.
10.	assertion	therefore specialisation of labour possible.
11.	example	different trades began like weaver etc.
12.	qualification	no major technical changes except in agricultural sphere.
13.	repetition	possibility of development depended on production of surplus food.

Sentence No.	Rhetorical Analysis	Discourse of Content
14.	repetition	so specialisation of labour possible.

SUMMARY: Agriculture had profound effects on human life.
Main function to produce surplus.
This permitted specialisation of labour.
This led to technical development in agricultural field.

PARAGRAPH IV		
15.	assertion	an important contribution to farming technique was made by
	link	traction plough (paragraph 2).
	example	plough drawn by oxen.
16.	evidence	arable land cultivated by one man increased in size.
17.	summing up	the result : (topic) domestic animals as source of power very important.
18.	assertion	another technique : irrigation.
19.	evidence	facilities for grain storage
	example	in Egypt prove this fact.
20 & 21	qualification	comparison with modern standards show the increase was relative and good for those times.

Sentence No.	Rhetorical Analysis	Discourse Function of Content
SUMMARY:	Important contribution to farming technique was traction plough. Another technique was irrigation.	
PARAGRAPH V		
22.	assertion	(topic) with better farming techniques, conditions ready for development of civilisation. (logical development)
23.	assertion	civilisation called the culture of cities.
24.	assertion	cities are gatherings of men not engaged in food production.
25 & 26	example	locations of cities and civilisations.
27.	repetition (link)	repetition of earlier assertion in paragraph 3 about importance of surplus food.
28.	elaboration	(reason for link) additional value of the surplus.
29.	assertion	the surplus is a basis for trade.
30.	evidence	certain commodities mined in one place exchanged for food.
31.	elaboration	different goods made that could be exchanged, e.g.

Sentence No.	Rhetorical Analysis	Discourse Function of Content
	example	flint, copper, made into axes, also ornaments etc.
32.	assertion	with growing size and complexity of communities, need for central administration.
33.	evidence	cities became centres for kings etc. who took tribute in form of food in return for government.
34.	summing up	all the above conditions led to origin of state.

SUMMARY: With better conditions, civilisation possible.
This associated with culture of cities which are gatherings of men not engaged on food production.
With surplus food, trade possible.
With complexity of communities, need for central administration resulting in state.

PARAGRAPH VI		
35.	assertion	(refers to main theme again) history of civilisation shows the development of science was in answer to man's needs.
36.	evidence	invention of the wheel used for help in transport and

Sentence No.	Rhetorical Analysis	Discourse Function of Content
		production of consumer goods.
37.	evidence	sailing ship invented and navigational methods.
38.	evidence	science of astronomy developed.
39	evidence	need for planning of crops so calendar invented.
40.	elaboration	more information about calendar.
41.	qualification	advantages of these inventions reserved for ruling class.

SUMMARY: Needs of man grow, and in answer new inventions made to satisfy them.
Evidence for this seen in: science of astronomy and navigation, transport etc., and a calendar for agriculture.

PARAGRAPH VII		
42.	assertion	with growth of cities, technical problems arose which had to be solved.
43.	evidence	one problem was supply of water and disposal of waste.
44.	example	city of Mohenjo-daro with evidence of town-planning.

SUMMARY: With cities, technical problems arose which had to be solved by town planning.

Sentence No.	Rhetorical Analysis	Discourse Function of Content
PARAGRAPH VIII		
45.	assertion	(reference to main theme again) civilisation creates wealth, technology and leisure which leads to development of arts.
46.	evidence	earliest times : with meat, artists painted cave walls.
47.	evidence	later architecture developed also.
	example	in Egypt musical scale developed.
48.	summing up	(topic and main theme) on basis of successful agriculture, main features of civilisation developed 5000 years before industrial revolution.

SUMMARY: With civilisation, not only wealth and technology, but leisure for arts also was possible.
With successful agriculture, the main features of civilisation were developed more than 5000 years ago.

3.A.2.

ANALYSIS OF TEXT : PASSAGE 3 A

Sentence No.	Rhetorical Analysis	Discourse Function of Content
PARAGRAPH I		
1.	assertion	topic : first cities appeared 5000 years ago in food-producing areas of Middle East
2.	example assertion	cities of Sumeria etc. had a number of distinguishing features
3.	evidence	(1) larger, more populated than surrounding villages
4.	evidence	(2) specialisation of labour produced complex hierarchy of social classes
5.	evidence	(3) the capital for cities came from villages providing surplus food
6.	evidence	(4) increased sophistication resulted in development of writing, counting etc.

SUMMARY: Origin of cities was in the Middle East 5000 years ago.
They had distinctive features that differentiated them from surrounding villages.
They were larger and more populated with a complex hierarchy of social classes
Money for development came from surplus food grown in villages.
This led to development of arts in cities.

Sentence No.	Rhetorical Analysis	Discourse Function of Content
PARAGRAPH II		
7.	link	(connection with Paragraph I) basis of European city lay in Middle East via Greece and Rome
8.	assertion	main theme : three main phases in growth of W. European city
9.	elaboration	(logical development) (1) Medieval phase from 11 century to 1500
10.	elaboration	(2) Renaissance and Baroque from 1500 to early 19 century
11.	elaboration	(3) early 19 century to modern period

SUMMARY: West European city developed in three phases: the Medieval, the Renaissance and Baroque, and the modern.

PARAGRAPH III		
12.	link assertion	(with paragraph 2) topic: every medieval city began with geographical or cultural centre.
13.	evidence	(logical development)

Sentence No.	Rhetorical Analysis	Discourse Function of Content
		this was often a permanent structure like a castle or church (cultural)
14.	evidence	a geographical centre was a market or trade route or river crossing
15.	assertion	(connection with theme) in urban geography, oldest part called nuclear settlement
16.	evidence	many small towns exist in Europe where this can be traced
17.	elaboration	this is difficult with the large modern city

SUMMARY: Every medieval city grew up around a geographical or cultural centre.
This was either a church or castle, or a trading centre.
The oldest part was called the nuclear settlement.

PARAGRAPH IV		
18.	assertion	(definition) a city essentially a group of dwellings where activities can be shared
19.	assertion	(more detail about main theme) all urban settlements

Sentence No.	Rhetorical Analysis	Discourse Function of Content
		must have certain basic requirements
20.	evidence	(logical development: list of requirements)
		(1) must be reasonably compact so as to be accessible
21.	evidence	(2) must have adequate space for traffic, offices and people
22.	evidence	(3) must be secure
23.	repetition	a summing up of earlier points: therefore earliest settlements built round castle
24 & 25.	elaboration	later they built wall round the town
26.	summing up	each town had become a fortress in Medieval period

SUMMARY: City essentially a group of dwellings that enable inhabitants to share enterprises.

Basic requirements are it must be: compact, it must have sufficient space and it must be secure.

Sentence No.	Rhetorical Analysis	Discourse Function of Content
PARAGRAPH V		
27.	assertion	decision to begin settlement depended on 2 factors: politico-cultural and economic
28.	evidence	(for politico-cultural) they built fortress or church
29.	elaboration	they preferred hillside or protection from three sides
30.	elaboration	the fortress or church dominated the country-side and had natural defences
31.	evidence	(for economic reasons) accessibility for trade
32.	elaboration	for this, flat land near waterway
33.	repetition & assertion	two functions often combined in single settlement
34.	example	Greek cities built with hill-top and low-lying area
35.	example	France, Belgium, Germany
36.	repetition	(reinforcement and cohesion) castle or (and) cathedral on hill-top and main urban community below

SUMMARY: Decision to build depended on two factors: politico-cultural and economic.

Sentence No.	Rhetorical Analysis	Discourse Function of Content
		For the former, a dominant site best with natural defences. For the latter trading considerations vital.
PARAGRAPH VI		
37.	link	(referring again to paragraph 2)
	assertion	same ideas were repeated in the Middle Ages.
38.	elaboration	the same features are recognizable
SUMMARY:		Recurring system of urban planning in the Middle Ages.
PARAGRAPH VII		
39.	link	(connection with paragraph 2)
	assertion	some cities grew up during Renaissance, but modern city rooted in Industrial Revolution
40.	elaboration	(more about theme) urbanisation earlier affected only a minority
41.	assertion	from 19 century on, 3 major economic factors led to growth of cities on a new scale
42.	evidence	(1) invention of powerful

Sentence No.	Rhetorical Analysis	Discourse Function of Content
		machines created factories and need for labour force
43.	evidence	(2) large-scale transport and concentration of industry and population possible because
44.	evidence	(3) agricultural revolution produced all the food necessary to sustain urban population

SUMMARY: The modern city is a post-industrial-revolution phenomena.
From early 19 century on, 3 factors contributed.
The large-scale invention of machines led to factories and the growth of a labour force.
Transport and communication were necessary for the concentration of industry and population.
The agricultural revolution provided the food necessary for the work force.

PARAGRAPH VIII		
45.	link	(connection) process of growth still linked to industry but
	assertion	growth of city also due to administrative and

Sentence No.	Rhetorical Analysis	Discourse Function of Content
		commercial needs
46.	evidence	earlier on, only 3% lived in small towns
47.	evidence	today there are 30%
48.	evidence	larger cities are attracting more people, still growing
49.	assertion	(more about theme) with urbanisation social problems emerging
	link	(connection with paragraph I, sentence 4) repetition of earlier information

SUMMARY: Rate of urbanisation also linked to administration and commerce.

Large cities are still increasing in size.

And with this growth social problems are being created.

S.8.1.

LIST OF IDEAS BASED ON PASSAGE 1

List of ideas from Passage 1 for use in the summary.

1. The earliest civilisations in the Middle East coincided with wild cereals.
2. This was followed by the domestication of animals for food.
3. Then animals were used for traction and transport.
4. Dairy farming began and wool was used for clothes.
5. Agriculture had the most profound effects on human life.
6. The main purpose of agriculture was to produce a surplus of food.
7. This made the specialisation of labour possible.
8. Which in turn led to technical developments, firstly in agriculture.
9. The first contribution to farming technique was the traction plough.
10. The second was the use of irrigation.
11. With better farming, the beginning of civilisation was the result.
12. This is seen as the culture of cities.
13. Cities are gatherings of men not engaged in agriculture.
14. The surplus of food permitted trade to take place.
15. The growing complexity of the community led to the need for a central administration.
16. This led to the origin of the state.
17. The history of science is related to human needs.
18. Evidence for this is seen in development of astronomy, the

wheel and sailing-ships.

19. Also the invention of the calendar for use in agriculture.

20. With the growth of cities, technical problems were created.

21. Town planning was needed to solve them.

22. Civilisation creates wealth.

23 It also creates technology.

24 With leisure time the development of the arts was possible.

25 The main features of civilisation were laid 5000 years ago.

3.8.2.

LIST OF IDEAS BASED ON PASSAGE 3 A

List of ideas from Passage 3 A for use in the summary.

1. The origin of cities was in the Middle East about 5000 years ago.
2. They had distinctive features that differentiated them from the surrounding villages.
3. They were larger and more populated with a complex hierarchy of social classes.
4. Money for their development came from surplus food grown in the villages.
5. This led to the development of arts in the cities.
6. The West European city developed in three phases: the Medieval, the Renaissance and Baroque, and the modern.
7. Every medieval city grew up round a geographical or cultural centre.
8. This was either a church or castle, or a trading centre.
9. The oldest part of the town was called the nuclear settlement.
10. The city is essentially a group of dwellings that enable inhabitants to share enterprises.
11. The basic requirements are: it must be compact,
12. it must be secure,
13. it must have sufficient space.
14. The decision to build depended on two factors: politico-cultural and economic.
15. For the former, a dominant site was best with natural defences.
16. For the latter, trading considerations were vital.

17. In the Middle Ages, a recurring system of urban planning evident.

18. The modern city is a post-industrial-revolution phenomena.

19. From the early 19 century on, three factors contributed.

20. The large scale invention of machines led to the growth of factories and a labour force.

21. Development of transport and communications made this concentration of industry and population possible.

22. And the agricultural revolution provided the food for the work force.

23. Rate of urbanisation also linked to administration and commerce.

24. This is proved by large cities which are becoming larger.

25. And with this growth, social problems are being created.

3.C.

ORGANIZATION OF TEXT

In addition to analyzing the text in terms of discourse function, we applied a technique described by Thomas in the same paper referred to earlier. The purpose of his analysis was to ascertain the 'meaning structure' of a text.

'Certain items have to be held in memory until a later unit adds to their meaning or other units make cross references to previous ones etc. Thus the meaning of a text is better represented by some display which shows how each item relates to other items within the paragraph.' (Page 4) He proceeded to do this by using a matrix to chart the connection between sentences in the text. This he later converted into a 'flow diagram'. We used the matrix idea which seemed to exemplify *via* a visual display, what we meant by the organization of a text. For us, our interest lay in seeing how different parts of the whole were inter-related in terms of ideas.

It was appreciation of this organization that we were trying to test in the questions added to Version 3 of the Multiple-Choice Test.

We used a cross 'X' to indicate which sentences were connected. The nature of the connection is not explored or analyzed, only the fact of the relationship. This provides us with a pattern of internal cohesion structure, and the analysis of the two passages used in the summary question shows how texts can be different in this respect.

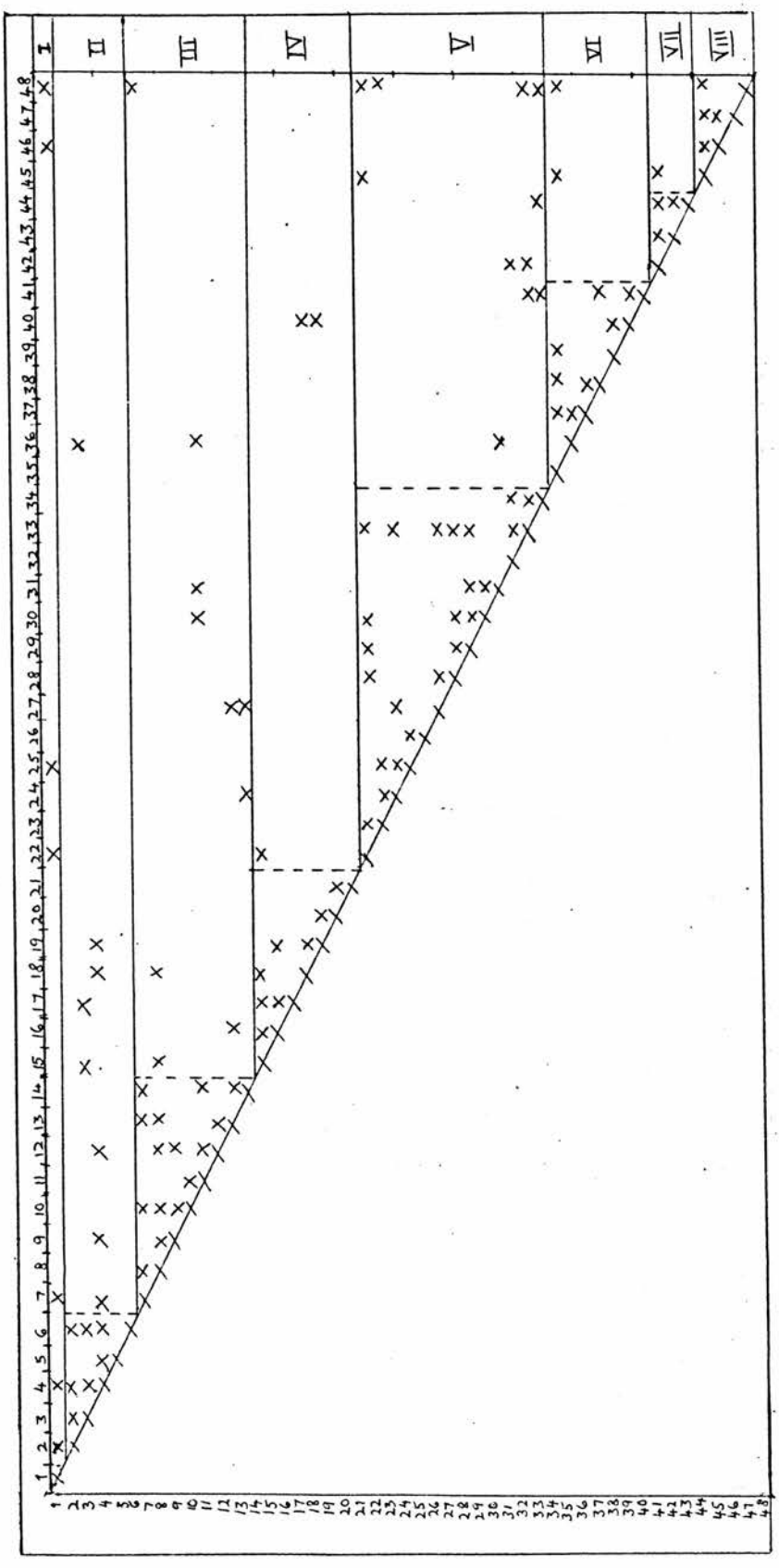
For any sentence, the vertical cross pattern shows links with earlier sentences, and the horizontal crosses indicate links to

subsequent sentences in the passage. By reference to the matrix, it is possible to identify all sentences that are linked by a common idea, or any other common denominator which is why we have labelled our matrix 'Organization of Text'.

In Passage 1, there is a much tighter structural link with constant references being made to earlier parts, and the beginning and end related making of the text a rounded whole. The different parts are seen to contribute to the central theme which is constantly kept to the forefront of attention by the use of devices like link, evidence, repetition etc.

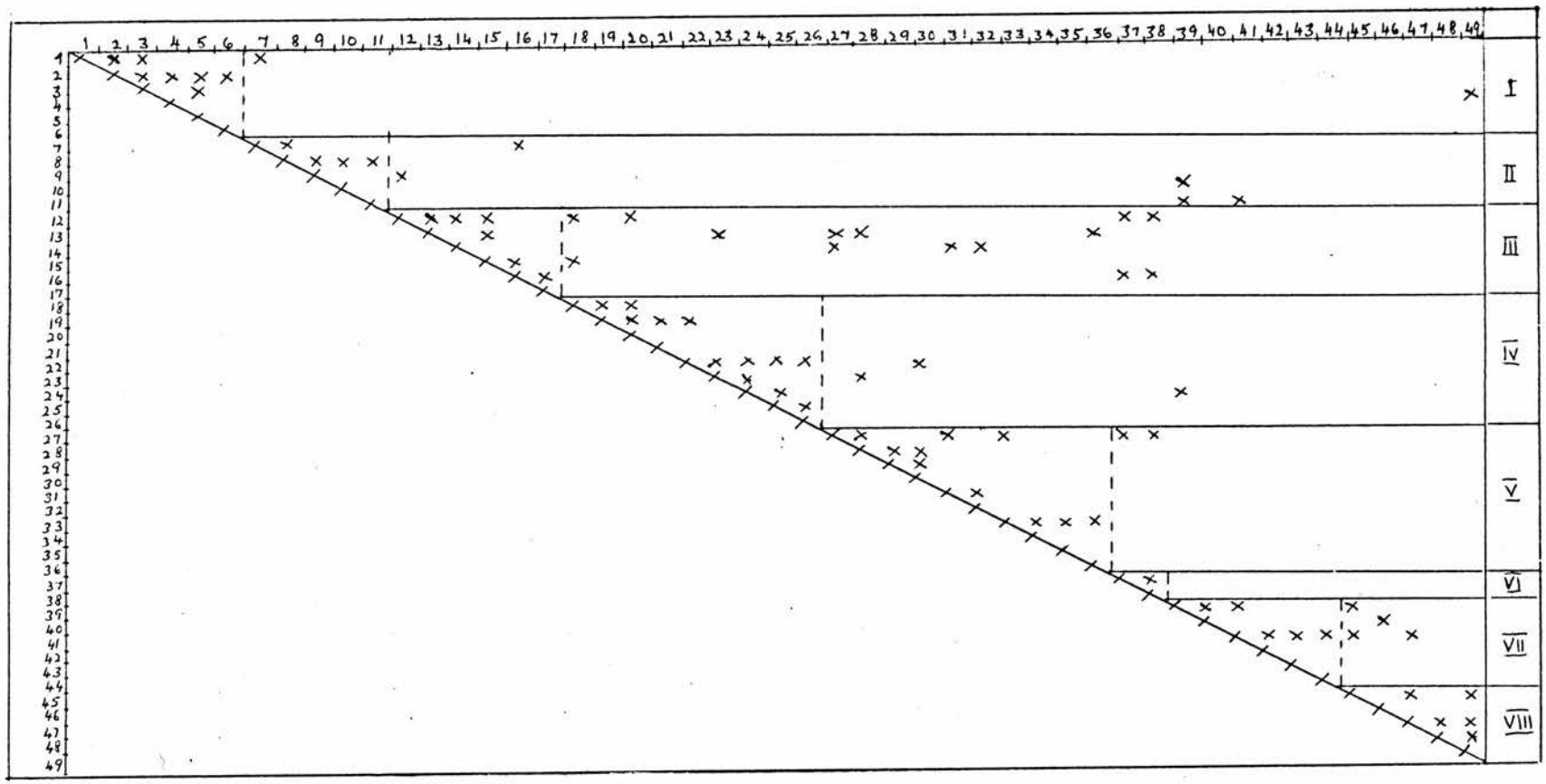
Passage 3 A is rather differently organized. It is a sequentially organized text where the consecutive units are logically related so that the content is presented more as a chain of ideas. Although the cohesion is indicated, the coherence is more tenuous and the NNS reader might have found it difficult to infer the super-ordinate concept of 'Urbanization' had the title not been provided.

We feel that this whole area is vastly underestimated as a source of difficulty for the foreign reader, and suggest that this is a field requiring further investigation.



3.c.1 ORGANIZATION OF TEXT : PASSAGE 1

Note: Numbers refer to sentences, and paragraphs are indicated by roman numerals in the extreme right hand column.



3.C.2. ORGANIZATION OF TEXT : PASSAGE 3 A

Note: Numbers refer to sentences, and paragraphs are indicated by roman numerals in the extreme right hand column.

APPENDIX FOUR

RESULTS

4.A. EXPERIMENT I : EFFECTS OF ORDERING

4.A.1. I (A) : INDIVIDUAL RESULTS

NO.	GROUP 1 (N = 25)				GROUP 2 (N=24)	
	CLOZE %		MULTIPLE-CHOICE	ELBA	MULTIPLE-CHOICE	ELBA
	EXACT	ACCEPTABLE	VERSION 1 (20 ITEMS)	%	VERSION 1 (20 ITEMS)	%
1	54	75	17	94	15	68
2	51	71	18	87	10	77
3	56	83	17	83	15	91
4	43	65	15	93	15	61
5	44	70	14	89	6	51
6	67	81	17	86	12	52
7	57	78	17	96	15	67
8	47	67	13	80	17	80
9	63	81	16	89	10	48
10	40	52	8	69	12	-
11	52	72	16	84	10	-
12	60	85	19	94	14	-
13	39	59	7	70	13	-
14	46	58	12	70	17	-
15	57	80	14	92	10	-
16	47	73	15	70	12	-
17	72	85	17	-	13	-
18	61	81	16	-	17	-
19	51	72	15	-	13	-
20	53	85	17	-	14	-
21	34	45	10	-	13	-
22	27	38	12	-	12	-
23	45	62	15	-	14	-
24	55	79	16	-	6	-
25	34	40	8	-	-	-

4.A.2. I (B) : INDIVIDUAL RESULTS

EGYPTIAN GROUP : N = 16

NO.	CLOZE %		MULTIPLE-CHOICE VERSION 3 (25 ITEMS)	INSTITUTE TEST %
	EXACT	ACCEPTABLE		
1	51	70	16	60
2	41	58	17	57
3	28	41	15	71
4	40	59	19	59
5	43	63	14	74
6	47	67	19	69
7	50	67	14	83
8	49	68	17	57
9	36	47	9	56
10	48	62	14	64
11	43	58	14	68
12	49	67	8	76
13	59	81	13	65
14	62	82	18	77
15	56	75	12	71
16	48	67	17	70

NOTE: Numbers 1 to 8 did the tests in the order listed i.e. cloze followed by multiple-choice. Numbers 9 to 16 did the tests in reverse order.

4.B. EXPERIMENT II PROCESS VERSUS PRODUCT

CLOZE VERSUS SUMMARY

4.B.1. GROUP A : INDIVIDUAL RESULTS N = 11

NO.	CLOZE PASSAGE 1 %		SUMMARY: PASSAGE 1 (25 ideas)	ELBA %
	EXACT	ACCEPTABLE		
1	11	14	6	44
2	16	21	7	60
3	21	25	6	54
4	21	27	4	54
5	22	35	8	56
6	23	38	9	47
7	24	27	8	44
8	28	34	6	40
9	32	44	7	59
10	32	45	10	54
11	37	54	14	52

NOTE: The summary was scored by analyzing the passage for the list of 25 ideas comprising the informational content. For details see Appendix 3.B.1.

4.8.2. GROUP B : INDIVIDUAL RESULTS N = 28

NO.	CLOZE PASSAGE I		SUMMARY PASSAGE 3A (25 ideas)	EXTERNAL CRITERION
	EXACT	ACCEPTABLE		
1	15	22	10	ELBA (%): 44
2	24	31	5	38
3	26	31	8	52
4	26	31	8	45
5	42	59	18	57.5
6	36	48	17	40
7	41	49	14	43
8	34	48	14	52
9	38	45	10	56
10	33	48	18	47
11	28	33	5	56
12	26	32	7	47
13	26	31	8	49
14	16	18	9	EPTB[1]: 34.9
15	18	21	13	25.7
16	21	26	8	35.6
17	19	21	7	25.6
18	21	28	8	22.3
19	18	25	9	33.8
20	44	52	15	36
21	30	43	14	34.7
22	25	31	8	35.4
23	29	40	8	36.2
24	36	54	14	38.4
25	25	32	11	34
26	25	34	12	34.1
27	52	75	16	52
28	24	33	10	-

NOTE: Details regarding list of ideas used for scoring the summary are contained in Appendix 3.B.2.

1. EPTB reports result in standard scores.

4.C. EXPERIMENT III : PROCESS VERSUS PRODUCT

ANALYSIS OF RESULTS : CLOZE TEST PASSAGE 1

4.C.1. BRITISH SAMPLE N = 190

	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
Items 1 - 10							
Correct	476	$\bar{x}$ = 7.56	353	$\bar{x}$ = 5.51	269	$\bar{x}$ = 4.27	1098
Incorrect	122	1.94	215	3.36	234	3.71	571
Omitted	32	.51	72	1.13	127	2.02	231
Items 11 - 20							
Correct	428	$\bar{x}$ = 6.79	302	$\bar{x}$ = 4.72	208	$\bar{x}$ = 3.30	938
Incorrect	175	2.78	255	3.98	272	4.32	702
Omitted	27	.43	83	1.30	150	2.38	260
Items 21 - 30							
Correct	349	$\bar{x}$ = 5.54	160	$\bar{x}$ = 2.5	94	$\bar{x}$ = 1.49	603
Incorrect	239	3.79	336	5.25	346	5.49	921
Omitted	42	.67	144	2.25	190	3.02	376
Items 31 - 40							
Correct	469	$\bar{x}$ = 7.44	339	$\bar{x}$ = 5.30	196	$\bar{x}$ = 3.11	1004
Incorrect	131	2.08	206	3.22	248	3.94	585
Omitted	30	.48	95	1.48	186	2.95	311
Items 41 - 50							
Correct	371	$\bar{x}$ = 5.89	215	$\bar{x}$ = 3.36	133	$\bar{x}$ = 2.11	719
Incorrect	203	3.22	253	3.95	304	4.83	760
Omitted	56	.89	172	2.69	193	3.06	421
Items 51 - 60							
Correct	429	$\bar{x}$ = 6.81	306	$\bar{x}$ = 4.78	153	$\bar{x}$ = 2.43	888
Incorrect	175	2.78	204	3.19	278	4.41	657
Omitted	26	.41	130	2.03	199	3.16	355

	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
Items 61 - 70							
Correct	359	$\bar{x}$ = 5.7	247	$\bar{x}$ = 3.86	121	$\bar{x}$ = 1.92	727
Incorrect	245	3.89	277	4.33	249	3.95	771
Omitted	26	.41	116	1.81	260	4.13	402
Items 71 - 80							
Correct	408	$\bar{x}$ = 6.48	294	$\bar{x}$ = 4.59	148	$\bar{x}$ = 2.35	850
Incorrect	193	3.06	252	3.94	218	3.46	663
Omitted	29	.46	94	1.47	264	4.19	387
Items 81 - 90							
Correct	521	$\bar{x}$ = 8.27	308	$\bar{x}$ = 4.81	120	$\bar{x}$ = 1.90	949
Incorrect	76	1.21	174	2.72	127	2.02	377
Omitted	33	.52	158	2.47	383	6.08	574
Items 91 - 100							
Correct	346	$\bar{x}$ = 5.49	166	$\bar{x}$ = 2.59	57	$\bar{x}$ = .90	569
Incorrect	205	3.25	231	3.61	137	2.17	573
Omitted	79	1.25	243	3.80	436	6.92	758
Total No.	6300		6400		6300		19000

4.C.1.1. ITEM ANALYSIS : BRITISH SAMPLE N = 190

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	the*	63	48	50	161	.85	.21	
2	pigs	24	14	6	44	.23	.29	
3	ass	42	27	21	90	.47	.33	✓
4	the*	43	25	17	85	.45	.41	✓✓
5	and*	60	40	25	125	.66	.56	✓✓
6	settlements	4	0	0	4	.02	.06	
7	in*	58	45	34	137	.72	.38	✓✓
8	of*	63	60	52	175	.92	.17	
9	years	61	54	47	162	.85	.22	
10	have*	58	40	17	115	.61	.65	✓✓
11	it*	44	42	42	128	.67	.03	
12	human	31	19	21	71	.37	.16	
13	and*	56	32	21	109	.57	.56	✓✓
14	it*	55	41	20	116	.61	.56	✓✓
15	activities	12	1	2	15	.08	.16	
16	rare	32	19	8	59	.31	.38	✓
17	the*	62	57	49	168	.88	.21	
18	as*	61	46	29	136	.72	.51	✓✓
19	among*	19	15	6	40	.21	.21	
20	became	56	30	10	96	.45	.73	✓
21	complex	28	14	2	44	.23	.41	
22	established	10	2	-	12	.06	.16	
23	flint	27	11	-	38	.2	.43	
24	technical	31	5	4	40	.21	.43	
25	the*	60	57	45	162	.85	.24	
26	place	40	8	6	54	.28	.54	✓
27	a*	45	14	8	67	.35	.59	✓✓
28	the*	58	39	20	117	.62	.60	✓✓
29	other	24	7	9	40	.21	.24	
30	contribution	26	3	0	29	.15	.41	
31	the*	50	27	20	97	.51	.48	✓✓
32	area	46	22	1	69	.36	.71	✓
33	of*	63	57	51	171	.90	.19	

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
34	fact	26	10	2	38	.20	.38	
35	was	61	51	28	140	.74	.52	✓
36	scale	49	31	12	92	.48	.59	✓
37	advances	47	26	9	82	.43	.60	✓
38	of*	14	5	1	20	.11	.21	
39	of*	61	59	39	159	.84	.35	✓✓
40	the*	52	51	33	136	.72	.30	✓✓
41	the*	48	20	6	74	.39	.67	✓✓
42	a*	32	6	3	41	.22	.46	
43	1950's	5	0	1	6	.03	.06	
44	kg	50	41	32	123	.65	.29	✓
45	practice	31	19	9	59	.31	.35	✓✓
46	invent	21	7	0	28	.15	.33	
47	called	20	3	1	24	.13	.30	
48	of*	56	31	23	110	.58	.52	✓✓
49	in*	56	46	29	131	.69	.43	✓✓
50	and*	52	42	29	123	.65	.37	✓✓
51	perhaps*	42	20	9	71	.39	.52	✓✓
52	arose	30	8	5	43	.23	.40	
53	an*	54	54	33	141	.74	.33	✓✓
54	workers	36	17	4	57	.30	.51	✓
55	the*	37	38	18	93	.49	.30	✓✓
56	it*	60	54	36	150	.79	.38	✓✓
57	such*	55	35	13	103	.54	.67	✓✓
58	are*	55	45	26	126	.66	.46	✓✓
59	transported	47	28	7	82	.43	.63	✓
60	either*	13	7	2	22	.12	.17	
61	finished	26	10	3	39	.21	.37	
62	other	26	18	8	52	.27	.29	✓
63	the*	56	45	27	128	.67	.46	✓✓
64	complex	28	17	9	54	.28	.30	✓
65	system	25	10	10	45	.24	.24	
66	cities	41	36	20	97	.51	.33	✓
67	trade	33	10	3	46	.24	.48	
68	to*	51	35	9	95	.50	.67	✓✓

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
69	took	31	20	7	58	.31	.38	✓
70	was	42	46	25	113	.59	.27	✓
71	and*	19	8	4	31	.16	.24	
72	ancient	37	26	12	75	.39	.40	✓
73	development	45	28	13	86	.45	.51	✓
74	and*	54	45	37	136	.72	.27	✓✓
75	used	52	46	20	118	.62	.51	✓
76	oversea	12	2	1	15	.08	.17	
77	and*	52	38	11	101	.53	.65	✓✓
78	methods	24	11	2	37	.19	.35	
79	the*	59	52	27	138	.73	.51	✓✓
80	to*	54	38	21	113	.59	.52	✓✓
81	and*	54	38	21	113	.59	.52	✓✓
82	year	62	50	31	143	.75	.49	✓
83	calendar	56	44	13	113	.59	.68	✓
84	astronomy	47	17	3	67	.35	.70	✓
85	a*	54	33	6	93	.49	.76	✓✓
86	a*	43	17	2	62	.33	.65	✓✓
87	it*	55	26	6	87	.46	.78	✓✓
88	of*	42	27	9	78	.41	.52	✓✓
89	cities	51	24	10	85	.45	.65	✓
90	one*	57	32	19	108	.57	.60	✓✓
91	and*	27	17	5	49	.26	.35	✓✓
92	of*	55	35	13	103	.54	.67	✓✓
93	by*	41	28	12	81	.43	.46	✓✓
94	and*	52	28	6	86	.45	.73	✓✓
95	leads	39	10	0	49	.26	.62	✓
96	earlier	9	3	1	13	.07	.13	
97	meat	22	3	0	25	.13	.35	
98	cave	24	3	1	28	.15	.37	
99	also*	24	4	1	29	.15	.37	
100	a*	53	35	18	106	.56	.56	✓✓
	TOTAL	4156	2690	1499	8345			

NOTES: 1. Items with an asterisk are functional items.

2. Items that are considered discriminating are indicated in the COMMENT column: one tick = content, double tick = functional.

4.C.1.2. INCORRECT AND OMITTED RESPONSES : BRITISH SAMPLE N = 190

	INCORRECT RESPONSES				OMITTED RESPONSES			
ITEM NO.	UPPER	MIDDLE	LOWER	TOTAL	UPPER	MIDDLE	LOWER	TOTAL
1	0	13	9	22	0	3	4	7
2	24	31	41	96	15	19	16	50
3	12	21	24	57	9	16	18	43
4	16	32	35	83	4	7	11	22
5	3	20	22	45	0	4	16	20
6	57	54	45	156	2	10	18	30
7	3	11	9	23	2	8	20	30
8	0	2	6	8	0	2	5	7
9	2	8	10	20	0	2	6	8
10	5	23	33	61	0	1	13	14
11	17	14	14	45	2	8	7	17
12	27	41	38	106	5	4	4	13
13	5	20	22	47	2	12	20	34
14	8	15	22	45	0	8	21	29
15	42	53	35	130	9	10	26	45
16	30	36	36	102	1	9	19	29
17	1	5	11	17	0	2	3	5
18	2	18	24	44	0	0	10	10
19	37	32	42	111	7	17	15	3
20	6	21	28	55	1	13	25	39
21	34	28	39	101	1	22	22	45
22	46	42	38	126	7	20	25	52
23	22	19	15	56	14	34	48	96
24	19	40	34	93	13	19	25	57
25	3	6	7	16	0	1	11	12
26	23	50	45	117	0	6	12	18
27	17	39	42	98	1	11	13	25
28	4	20	33	57	1	5	10	16
29	34	39	39	112	5	18	15	38
30	37	53	54	144	0	8	9	17
31	9	17	24	50	4	20	19	43
32	16	37	34	87	1	5	28	34
33	0	4	9	13	0	3	3	6

ITEM NO.	INCORRECT RESPONSES				OMITTED RESPONSES			
	UPPER	MIDDLE	LOWER	TOTAL	UPPER	MIDDLE	LOWER	TOTAL
34	25	31	28	84	12	23	33	68
35	1	6	9	16	1	7	26	34
36	14	32	41	87	0	1	10	11
27	15	36	44	95	1	2	10	13
38	43	34	38	115	6	25	24	55
39	2	3	12	17	0	2	12	14
40	6	6	9	21	5	7	21	33
41	11	19	26	56	4	25	31	60
42	26	43	39	108	5	15	21	41
43	33	20	20	73	25	44	42	111
44	9	7	10	26	4	16	21	41
45	32	35	35	102	0	10	19	29
46	38	43	46	127	4	14	17	35
47	31	32	45	108	12	29	17	58
48	7	23	27	57	0	10	13	23
49	7	17	28	52	0	1	6	7
50	9	14	28	51	2	8	6	16
51	16	25	32	73	5	19	22	46
52	32	45	40	117	1	11	18	30
53	8	5	24	37	1	5	6	12
54	22	20	30	72	5	27	29	61
55	25	23	26	74	1	3	19	23
56	2	9	15	26	1	1	12	14
57	5	16	20	41	3	13	30	46
58	8	10	26	44	0	9	11	20
59	13	21	33	67	3	15	23	41
60	44	30	32	106	6	27	29	62
61	36	39	35	110	1	15	25	41
62	32	27	29	88	5	19	26	50
63	6	13	9	28	1	6	27	34
64	34	41	34	109	1	6	20	27
65	36	43	31	110	2	11	22	35
66	18	21	13	52	4	7	30	41
67	25	24	14	63	5	30	46	81
68	10	20	28	58	2	9	26	37

ITEM NO.	INCORRECT RESPONSES				OMITTED RESPONSES			
	UPPER	MIDDLE	LOWER	TOTAL	UPPER	MIDDLE	LOWER	TOTAL
69	29	33	38	100	3	11	18	32
70	19	16	18	53	2	2	20	24
71	35	36	34	105	9	20	25	54
72	23	36	38	97	3	2	13	18
73	17	28	23	68	1	8	27	36
74	9	17	15	41	0	2	11	13
75	9	12	16	37	2	6	27	35
76	47	52	32	131	4	10	30	44
77	10	12	16	38	1	14	36	51
78	32	35	20	87	7	18	41	66
79	4	7	10	21	0	5	26	31
80	7	17	14	38	2	9	28	39
81	2	5	5	12	7	21	37	65
82	1	4	6	11	0	10	26	36
83	4	10	11	25	3	10	39	52
84	12	27	17	56	4	20	43	67
85	8	21	15	44	1	10	42	53
86	18	29	20	67	2	18	41	61
87	6	24	11	41	2	14	46	62
88	17	19	15	51	4	18	39	61
89	4	17	15	36	8	23	38	69
90	4	18	12	34	2	14	32	48
91	34	29	23	86	2	18	35	55
92	5	12	11	28	3	17	39	59
93	16	20	9	45	6	16	42	64
94	6	10	11	27	5	26	46	77
95	18	24	13	55	6	30	50	86
96	44	40	20	104	10	21	42	73
97	25	28	12	65	16	33	51	100
98	32	37	14	83	7	24	48	79
99	21	26	17	64	18	34	45	97
100	4	5	7	16	6	24	38	68
TOTAL	1764	2403	2413	6580	380	1307	2388	4075

4.C.2. ITEM ANALYSIS OF CLOZE TEST PASSAGE 1

NUREMBERG SAMPLE N = 53

ITEM NO.	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	18	15	16	49	.92	.11	
2	7	4	8	19	.36	-.05	
3	13	7	7	27	.51	.33	✓
4	14	13	9	36	.68	.28	✓✓
5	17	15	9	41	.77	.44	✓✓
6	-	-	-	-	-	-	
7	18	17	9	44	.83	. 5	✓✓
8	18	17	18	53	1.00	0	
9	18	17	17	52	.98	.05	
10	18	17	14	49	.92	.22	
11	15	12	15	42	.79	0	
12	15	12	12	39	.74	.17	
13	15	10	12	37	.70	.17	
14	17	15	10	42	.79	.39	✓✓
15	2	-	-	2	.04	.11	
16	6	2	4	12	.23	.11	
17	18	16	15	49	.92	.17	
18	16	16	15	47	.89	.05	
19	7	4	-	11	.21	.39	
20	16	13	12	41	.77	.22	
21	7	5	3	15	.28	.22	
22	-	1	-	1	.02	-	
23	6	3	2	11	.21	.22	
24	8	1	1	10	.19	.39	
25	17	16	14	47	.89	.17	
26	5	1	3	9	.17	.11	
27	18	15	10	43	.81	.44	✓✓
28	15	15	16	46	.87	-.05	
29	4	10	9	23	.43	-.27	
30	3	1	-	4	.08	.17	
31	14	15	11	40	.75	.17	
32	10	10	10	30	.57	0	

ITEM NO.	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
33	18	16	18	52	.98	0	
34	6	4	4	14	.26	.11	
35	18	16	17	51	.96	.05	
36	10	5	4	19	.36	.33	✓
37	9	7	5	21	.40	.22	
38	6	2	-	8	.15	.33	
39	17	15	15	47	.89	.11	
40	18	11	14	43	.81	.22	
41	13	13	5	31	.58	.44	✓✓
42	11	10	6	27	.51	.28	✓✓
43	-	-	-	-	-	-	
44	16	17	11	44	.83	.28	✓
45	7	6	2	15	.28	.28	✓
46	7	8	3	18	.34	.22	
47	3	-	3	6	.11	0	
48	16	13	10	39	.74	.33	✓✓
49	18	15	17	50	.94	.05	
50	18	15	15	48	.91	.17	
51	15	6	5	26	.49	.56	✓✓
52	8	3	3	14	.26	.28	✓
53	18	15	13	46	.87	.28	
54	10	8	5	23	.43	.28	✓
55	16	13	6	35	.66	.56	✓✓
56	18	14	15	47	.89	.17	
57	14	12	10	36	.68	.22	
58	13	9	12	34	.64	.05	
59	12	7	5	24	.45	.39	✓
60	2	2	3	7	.13	-.05	
61	2	2	-	4	.08	.11	
62	7	5	4	16	.30	.17	
63	18	16	12	46	.87	.33	
64	3	3	1	7	.13	.11	
65	4	2	-	6	.11	.22	
66	11	8	6	25	.47	.28	✓
67	5	4	7	16	.30	-.11	
68	18	16	13	47	.89	.28	

ITEM NO.	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
69	9	3	6	18	.34	.17	
70	17	15	13	45	.85	.22	
71	5	7	2	14	.26	.17	
72	9	6	3	18	.34	.33	✓
73	9	3	6	18	.34	.17	
74	16	16	12	44	.83	.22	
75	15	11	5	31	.58	.56	✓
76	5	1	1	7	.13	.22	
77	18	16	12	46	.87	.33	
78	3	2	1	6	.11	.11	
79	18	14	13	45	.85	.28	✓✓
80	16	10	10	36	.68	.33	✓✓
81	18	15	12	45	.85	.33	✓✓
82	18	15	14	47	.89	.22	
83	14	9	8	31	.58	.33	✓
84	10	13	10	33	.62	0	
85	17	16	13	46	.87	.22	
86	16	11	7	34	.64	.5	✓✓
87	15	15	9	39	.74	.33	✓✓
88	10	9	4	23	.43	.33	✓✓
89	12	9	5	26	.49	.39	✓
90	17	16	8	41	.77	.5	✓✓
91	7	2	2	11	.21	.28	
92	18	16	10	44	.83	.44	✓✓
93	10	12	6	28	.53	.22	
94	16	13	6	35	.66	.56	✓✓
95	15	9	2	26	.49	.72	✓
96	1	3	1	5	.09	0	
97	9	5	2	16	.30	.39	✓
98	11	4	2	17	.32	.5	✓
99	6	-	-	6	.11	.33	
100	18	14	7	39	.74	.61	✓✓
TOTAL	1178	943	762	2883			

NOTE: Items that are considered discriminating are indicated in the COMMENT column: one tick = content, double tick = functional

4.C.3. EXPERIMENT III : INDIVIDUAL RESULTS

BRITISH SAMPLE

For reasons of convenience, all results are reported together to avoid duplication. The country of origin has been included for information only. As already reported, it was at one time planned to use this information, though this idea was subsequently discarded.

NUMBER	CLOZE %		MULTIPLE-CHOICE	EXTERNAL CRITERION	COUNTRY
	EXACT	ACCEPTABLE			
			VERSION 1.		
1	54	75	17	ELBA: 94	INDIA
2	51	71	18	87	INDIA
3	56	83	17	83	EGYPT
4	43	65	15	93	CHILE
5	44	70	14	89	BRAZIL
6	67	81	17	86	ITALY
7	57	78	17	96	CHILE
8	47	67	13	80	HONG-KONG
9	63	81	16	89	INDIA
10	40	52	8	69	INDONESIA
11	52	72	16	84	PAKISTAN
12	60	85	19	94	FINLAND
13	39	59	7	70	SYRIA
14	46	58	12	70	INDONESIA
15	57	80	14	92	MALAYSIA
16	47	73	15	70	HONG-KONG
17	72	85	17	-	SPAIN
18	61	81	16	-	KENYA
19	51	72	15	-	MEXICO
20	53	85	17	-	NORWAY
21	34	45	10	-	SPAIN

NUMBER	CLOZE %		MULTIPLE-CHOICE	EXTERNAL CRITERION	COUNTRY
	EXACT	ACCEPTABLE			
22	27	38	12	-	HONG-KONG
23	45	62	15	-	SUDAN
24	55	79	16	-	GERMANY
25	34	40	8	-	SPAIN
26	24	31	-	ELBA: 38	GREECE
27	15	22	-	44	IRAN
28	8	10	-	43	JORDAN
29	28	33	-	56	MALAYSIA
30	33	48	-	47	SAUDI ARABIA
31	38	45	-	56	HONG KONG
32	34	48	-	52	YUGOSLAVIA
33	41	49	-	43	COLOMBIA
34	26	31	-	52	JAPAN
35	36	48	-	40	INDIA
36	22	35	-	56	LEBANON
37	23	38	-	47	KUWAIT
38	32	45	-	54	IRAQ
39	32	44	-	59	TAIWAN
40	28	34	-	40	VENEZUELA
41	43	60	-	76	JAPAN
42	54	71	-	94	MEXICO
43	52	71	-	84	IRAQ
44	44	58	-	75	HONG-KONG
45	37	56	-	52	SPAIN
46	42	59	-	57.5	GREECE
47	22	27	-	54	GREECE
48	21	25	-	54	YUGOSLAVIA
49	16	22	-	60	JORDAN
50	11	14	-	44	HONG-KONG
51	24	27	-	42	INDONESIA
52	26	32	-	47	HONG-KONG
53	26	31	-	49	MALAYSIA
54	26	31	-	45	IRAQ
55	24	33	-	-	GERMANY

NUMBER	CLOZE %		MULTIPLE-CHOICE	EXTERNAL CRITERION	COUNTRY
	EXACT	ACCEPTABLE			
56	36	51	-	-	VENEZUELA
57	25	35	-	EPTB: 34.7	INDONESIA
58	25	32	-	34	MEXICO
59	25	31	-	35.4	INDONESIA
60	21	28	-	22.3	COLOMBIA
61	46	65	-	-	VENEZUELA
62	36	54	-	38.4	MEXICO
63	40	52	-	-	VENEZUELA
64	52	75	-	52	KENYA
65	43	65	-	-	VENEZUELA
66	47	69	-	-	VENEZUELA
67	42	57	-	-	ARGENTINE
68	38	58	-	-	ARGENTINE
			VERSION 2		
69	16	21	13	35.2	INDONESIA
70	35	44	14	39.1	THAILAND
71	32	48	15	42.2	THAILAND
72	24	39	15	32.6	INDONESIA
73	24	29	13	35.3	INDONESIA
74	26	38	13	40.2	THAILAND
75	27	36	13	39.9	INDONESIA
76	46	68	13	42.5	MEXICO
77	27	35	9	38.5	THAILAND
78	27	39	10	38.5	THAILAND
79	33	38	12	40.9	INDONESIA
80	27	31	11	27.1	NEPAL
81	37	48	10	40.8	THAILAND
82	34	42	11	31.7	ZAIRE
83	40	49	12	31.9	MEXICO
84	27	31	10	36	THAILAND
85	45	64	12	42.3	THAILAND
86	14	19	9	32.9	YEMEN
87	24	30	9	36.4	NIGERIA

NUMBER	CLOZE %		MULTIPLE-CHOICE	EXTERNAL CRITERION		COUNTRY
	EXACT	ACCEPTABLE				
88	15	20	9	EPTB:	32.6	INDONESIA
89	21	30	8		32	NEPAL
90	26	29	5		32.1	NEPAL
91	31	40	7		35.5	THAILAND
92	15	20	6		32.3	INDONESIA
93	25	31	7		26.5	YEMEN
94	14	17	5		33.4	INDONESIA
95	10	17	8		32.5	YEMEN
96	15	18	5		32.6	MEXICO
97	22	26	8		31.4	SAUDI ARABIA
98	29	38	5		35.1	COLOMBIA
99	29	40	8		35.6	THAILAND
100	27	38	8		34.5	THAILAND
101	20	25	8		32.3	INDONESIA
102	37	47	8		33.7	VENEZUELA
103	24	33	13		-	JAPAN
104	47	61	17		-	JAPAN
105	42	54	13		-	BRAZIL
106	41	50	14		-	MEXICO
107	44	55	14		-	JAPAN
108	53	65	14		-	MOROCCO
109	31	42	15	LANG.T.	70	JAPAN
110	21	26	16		72	BRAZIL
111	36	52	16		72	BRAZIL
112	21	25	13		74	CHILE
113	19	26	14		74	MEXICO
114	47	57	13		76	MEXICO
115	25	34	14		64	MEXICO
116	37	44	17		85	MEXICO
117	48	61	15		82	GREECE
118	29	38	13		78	MEXICO
119	40	49	15		79	MOROCCO
120	55	63	13		48	MAHE
121	41	53	16		60	COLOMBIA

NUMBER	CLOZE %		MULTIPLE-CHOICE	EXTERNAL CRITERION	COUNTRY
	EXACT	ACCEPTABLE			
122	16	20	10	-	MEXICO
123	14	18	10	-	MOROCCO
124	35	47	11	-	BRAZIL
125	46	56	11	-	COSTA RICA
126	39	49	12	-	BRAZIL
127	20	21	12	LANG.T. 56	BRAZIL
128	25	29	11	68	MEXICO
129	27	39	10	60	MEXICO
130	21	24	11	62	MOROCCO
131	26	38	12	63	CHILE
132	33	45	10	98	LIBYA
133	24	32	11	85	THAILAND
134	16	22	10	76	MEXICO
135	13	14	10	46	SAUDI ARABIA
136	29	35	12	60	CHILE
137	15	21	10	58	CHILE
138	29	37	9	-	MOROCCO
139	11	14	9	50	MEXICO
140	26	32	9	76	IRAN
141	16	19	9	46	VENEZUELA
142	10	13	7	45	MEXICO
143	46	62	8	83	MOROCCO
144	39	50	7	79	LIBYA
145	20	24	7	79	ARMENIAN
146	6	9	8	48	COLOMBIA
147	10	14	8	56	CHILE
148	36	49	3	-	NEPAL
149	24	33	7	-	JAPAN
150	18	19	5	-	LIBYA
151	21	30	5	-	LIBYA
152	14	20	3	LANG.T. 60	LIBYA
153	23	27	4	-	NAMBIA
154	30	38	7	-	NAMBIA

NUMBER	CLOZE %		MULTIPLE-CHOICE	EXTERNAL CRITERION	COUNTRY
	EXACT	ACCEPTABLE			
155	22	26	8	-	AFGHANISTAN
156	22	22	7	-	IRAN
157	33	40	6	-	MEXICO
158	16	18	13	EPTB: 34.9	INDONESIA
159	18	21	5	25.7	CHILE
160	21	26	15	35.6	INDONESIA
161	19	21	8	25.6	KUWAIT
162	29	42	13	36.2	INDONESIA
163	18	25	11	33.8	INDONESIA
164	45	54	16	36	VENEZUELA
165	30	43	14	34.7	NEPAL
166	49	68	VERSION 3 13	45.2	ZAIRE
167	50	61	14	41.3	ZAIRE
168	39	54	8	40.3	ZAIRE
169	27	36	10	-	MEXICO
170	18	20	7	-	NEPAL
171	26	37	12	-	INDONESIA
172	39	54	14	-	VENEZUELA
173	39	47	16	-	JAPAN
174	28	37	11	-	BRAZIL
175	41	51	12	-	/NGAMBAYE/
176	47	64	18	-	ALGERIA
177	57	79	21	-	FINLAND
178	53	70	15	-	GREECE
179	34	55	11	-	INDIA
180	28	40	11	-	ALGERIA
181	48	61	14	-	GAMBIA
182	39	48	14	-	JAPAN
183	34	42	14	-	JAPAN
184	39	57	17	-	MEXICO
185	31	39	10	-	IRAN
186	41	62	16	-	GERMANY
187	54	75	17	-	GERMANY
188	57	70	-	-	GERMANY
189	52	69	-	-	GERMANY
190	36	51	-	-	FRANCE

Additional subjects who sat only the Multiple-Choice Version 1.

NUMBER	MULTIPLE-CHOICE : VERSION 1	EXTERNAL CRITERION	COUNTRY
1	15	ELBA: 68	JAPAN
2	10	77	EGYPT
3	15	91	PAKISTAN
4	15	61	NEPAL
5	6	51	BRAZIL
6	12	52	INDIA
7	15	67	INDIA
8	17	80	INDIA
9	10	48	SRI LANKA
10	12	-	NIGERIA
11	9	-	SINGAPORE
12	14	-	KENYA
13	13	-	NIGERIA
14	17	-	RHODESIA
15	10	-	JAPAN
16	12	-	INDIA
17	13	-	HONG KONG
18	17	-	INDIA
19	13	-	GHANA
20	14	-	NORWAY
21	13	-	FRANCE
22	12	-	FRANCE
23	14	-	FRANCE
24	6	-	SWEDEN
25	6	-	SAUDI ARABIA
26	10	-	KUWAIT
27	8	-	TURKEY
28	7	-	OMAR
29	8	-	SAUDI ARABIA
30	13	-	TURKEY
31	12	-	TURKEY

NUMBER	MULTIPLE-CHOICE : VERSION 1	EXTERNAL CRITERION	COUNTRY
32	11	-	KUWAIT
33	6	-	SAUDI ARABIA
34	5	-	YEMEN
35	11	-	SAUDI ARABIA

NOTES: 1. For the external criterion, ELBA and LANG.T. scores are out of 100%. as are the Cloze results. EPTB reports standard scores.

2. Multiple-Choice Versions 1 and 2 have 20 items each, while Version 3 has 25.

4.C.4. EXPERIMENT III : INDIVIDUAL RESULTS

OVERSEAS SAMPLE: NUREMBERG UNIVERSITY N = 53

NUMBER	CLOZE %		MULTIPLE-CHOICE VERSION 2 (20 questions)	EXTERNAL CRITERION
	EXACT	ACCEPTABLE		
1	39	52	13	46.3
2	38	53	12	45
3	40	54	12	45.8
4	39	54	13	42.5
5	44	54	9	40
6	45	55	8	42.5
7	40	56	9	46.7
8	39	56	12	42.5
9	41	52	9	50.4
10	41	48	13	27.5
11	29	35	7	46.7
12	39	48	12	36.7
13	21	28	6	47.9
14	31	46	10	39.6
15	40	48	12	49.2
16	32	45	11	33.8
17	37	48	13	45.8
18	33	45	13	54.6
19	34	42	6	39.6
20	29	35	13	-
21	45	57	14	49.2
22	46	57	14	46.3
23	46	58	12	54.6
24	49	59	10	44.6
25	42	60	13	-
26	49	61	11	49.2
27	47	61	7	40.8
28	43	62	9	49.2

NUMBER	CLOZE %		MULTIPLE-CHOICE VERSION 2 (20 questions)	EXTERNAL CRITERION
	EXACT	ACCEPTABLE		
29	44	62	11	44.2
30	52	64	13	36.3
31	51	64	15	49.2
32	49	64	14	45
33	46	64	14	40.4
34	47	65	14	43.8
35	45	66	16	40
36	48	66	11	42.9
37	52	67	12	53.3
38	44	70	16	50.4
39	56	74	16	55.4
40	40	51	-	42.1
41	44	61	-	46.3
42	45	61	-	47.9
43	47	68	-	52.9
44	54	78	-	-
45	42	57	-	44.8
46	45	58	-	44.2
47	31	44	-	-
48	41	51	-	49.6
49	30	37	-	-
50	26	37	-	-
51	38	47	-	39.1
52	26	30	-	-
53	37	48	-	23.8

NOTE: The external criterion used was a language test administered by the English Department, Nuremberg University.

4.C.4. EXPERIMENT III : INDIVIDUAL RESULTS

OVERSEAS SAMPLE : KREFELD N = 16

NUMBER	CLOZE %		MULTIPLE-CHOICE VERSION 2 (20 questions)	EXTERNAL CRITERION
	EXACT	ACCEPTABLE		
1	56	65	14	-
2	53	66	10	-
3	47	60	8	-
4	44	55	10	-
5	54	65	9	-
6	51	68	12	-
7	56	66	14	-
8	59	72	11	-
9	46	57	12	-
10	55	68	15	-
11	57	69	14	-
12	55	69	14	-
13	51	66	13	-
14	55	68	11	-
15	46	54	8	-
16	60	71	10	-

4.D. EXPERIMENT IV : LENGTH AS VARIABLE

4.D.1. COMBINED ANALYSIS OF CLOZE DELETIONS : CONTENT VERSUS FUNCTIONAL ITEMS

ITEMS	P1	P2A	P3A	P4A	P5A
1 to 10	6F + 4C	5F + 5C	2F + 8C	4F + 6C	5F + 5C
11 to 20	6F + 4C	7F + 3C	3F + 7C	7F + 3C	5F + 5C
21 to 30	3F + 7C	3F + 7C	3F + 7C	2F + 8C	7F + 3C
31 to 40	5F + 5C	5F + 5C	3F + 7C	6F + 4C	6F + 4C
41 to 50	5F + 5C	6F + 4C	3F + 7C	5F + 5C	2F + 8C
51 to 60	7F + 3C	6F + 4C	3F + 7C	6F + 4C	5F + 5C
61 to 70	2F + 8C	1F + 9C	7F + 3C	4F + 6C	4F + 6C
71 to 80	5F + 5C	5F + 5C	5F + 5C	4F + 6C	4F + 6C
81 to 90	6F + 4C	7F + 3C	6F + 4C	9F + 1C	3F + 7C
91 to 100	6F + 4C	4F + 6C	2F + 8C	5F + 5C	5F + 5C
101 to 110				2F + 8C	
111 to 115				1F + 4C	
TOTAL	51F +49C	49F +51C	37F +63C	55F +60C	46F +54C

F = Functional Items

C = Content Items

4.D.2. COMBINED ANALYSIS OF CLOZE RESPONSES

PASSAGE 1 (British sample) N = 190			
Total Correct[1]	8345	Mean	43.92
Incorrect	6580		34.63
Omitted	4075		21.45

PASSAGE 2A N = 49				PASSAGE 2B N = 49			
Total Correct	1984	Mean	40.49	Total Correct	330	Mean	6.73
Incorrect	2078		42.41	Incorrect	688		14.04
Omitted	838		17.10	Omitted	207		4.22
PASSAGE 3A N = 36				PASSAGE 3B N = 36			
Total Correct	1351	Mean	37.53	Total Correct	360	Mean	10
Incorrect	1448		40.22	Incorrect	379		10.53
Omitted	801		22.25	Omitted	161		4.47
PASSAGE 4A N = 16				PASSAGE 4B N = 16			
Total Correct	1207	Mean	75.44	Total Correct	250	Mean	15.63
Incorrect	509		31.81	Incorrect	137		8.56
Omitted	124		7.75	Omitted	13		.81
PASSAGE 5A N = 16				PASSAGE 5B N = 16			
Total Correct	738	Mean	46.13	Total Correct	148	Mean	9.25
Incorrect	672		42	Incorrect	162		10.13
Omitted	190		11.88	Omitted	90		5.63

NOTE: 1. By acceptable word method.

4.0.2.1. RESULTS OF CLOZE TEST PASSAGE 2A N = 49

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
1 - 10							
Correct	92	$\bar{x}$ = 5.75	66	$\bar{x}$ = 3.88	39	$\bar{x}$ = 2.44	197
Incorrect	63	3.94	85	5	87	5.44	235
Omitted	5	.31	19	1.12	34	2.12	58
11 - 20							
Correct	99	$\bar{x}$ = 6.19	87	$\bar{x}$ = 5.12	58	$\bar{x}$ = 3.63	244
Incorrect	55	3.44	68	4	84	5.25	207
Omitted	6	.38	15	.88	18	1.12	39
21 - 30							
Correct	95	$\bar{x}$ = 5.94	85	$\bar{x}$ = 5	49	$\bar{x}$ = 3.06	229
Incorrect	51	3.19	57	3.35	82	5.13	190
Omitted	14	.88	28	1.65	29	1.81	71
31 - 40							
Correct	96	$\bar{x}$ = 6	78	$\bar{x}$ = 4.59	53	$\bar{x}$ = 3.31	227
Incorrect	58	3.63	78	4.59	82	5.13	218
Omitted	6	.38	14	.82	25	1.56	45
41 - 50							
Correct	114	$\bar{x}$ = 7.13	113	$\bar{x}$ = 6.64	71	$\bar{x}$ = 4.44	298
Incorrect	36	2.25	48	2.82	65	4.06	149
Omitted	10	.63	9	.53	24	1.5	43
51 - 60							
Correct	83	$\bar{x}$ = 5.19	71	$\bar{x}$ = 4.18	43	$\bar{x}$ = 2.69	197
Incorrect	70	4.38	74	4.35	76	4.75	220
Omitted	7	.44	25	1.47	41	2.56	73
61 - 70							
Correct	57	$\bar{x}$ = 3.56	25	$\bar{x}$ = 1.47	17	$\bar{x}$ = 1.06	99
Incorrect	97	6.06	83	4.88	60	3.75	240
Omitted	6	.38	62	3.65	83	5.19	151

ITEMS	UPPER THIRD	MIDDLE THIRD	LOWER THIRD	TOTAL
71 - 80				
Correct	69 $\bar{x}$ = 4.31	42 $\bar{x}$ = 2.47	27 $\bar{x}$ = 1.69	138
Incorrect	82 5.13	81 4.77	72 4.5	235
Omitted	9 .56	47 2.76	61 3.81	117
81 - 90				
Correct	85 $\bar{x}$ = 5.31	53 $\bar{x}$ = 3.12	30 $\bar{x}$ = 1.88	168
Incorrect	63 3.94	66 3.88	77 4.81	206
Omitted	12 .75	51 3	53 3.32	116
91 - 100				
Correct	87 $\bar{x}$ = 5.44	56 $\bar{x}$ = 3.29	44 $\bar{x}$ = 2.75	187
Incorrect	64 4	59 3.47	55 3.44	178
Omitted	9 .56	55 3.24	61 3.81	125
TOTAL	1600	1700	1600	4900

4.D.2.2. RESULTS OF CLOZE TEST PASSAGE 3A N = 36

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
1 - 10							
Correct	66	$\bar{x}$ = 5.5	44	$\bar{x}$ = 3.67	13	$\bar{x}$ = 1.08	123
Incorrect	45	3.75	60	5	48	4	153
Omitted	9	.75	16	1.33	59	4.92	84
11 - 20							
Correct	60	$\bar{x}$ = 5	57	$\bar{x}$ = 4.75	45	$\bar{x}$ = 3.75	162
Incorrect	48	4	54	4.5	51	4.25	153
Omitted	12	1	9	.75	24	2	45
21 - 30							
Correct	72	$\bar{x}$ = 6	47	$\bar{x}$ = 3.92	30	$\bar{x}$ = 2.5	149
Incorrect	40	3.33	50	4.17	56	4.67	146
Omitted	8	.67	23	1.92	34	2.83	65
31 - 40							
Correct	69	$\bar{x}$ = 5.75	43	$\bar{x}$ = 3.58	27	$\bar{x}$ = 2.25	139
Incorrect	47	3.92	54	4.5	41	3.42	142
Omitted	4	.33	23	1.92	52	4.33	79
41 - 50							
Correct	76	$\bar{x}$ = 6.33	54	$\bar{x}$ = 4.5	25	$\bar{x}$ = 2.08	155
Incorrect	35	2.92	58	4.83	51	4.25	144
Omitted	9	.75	8	.67	44	3.67	61
51 - 60							
Correct	61	$\bar{x}$ = 5.08	37	$\bar{x}$ = 3.08	11	$\bar{x}$ = .92	109
Incorrect	52	4.33	69	5.75	67	5.58	188
Omitted	7	.58	14	1.17	42	2.5	63
61 - 70							
Correct	66	$\bar{x}$ = 5.5	48	$\bar{x}$ = 4	21	$\bar{x}$ = 1.75	135
Incorrect	46	3.83	59	4.92	43	3.58	148
Omitted	8	.67	13	1.08	56	4.67	77

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
71 - 80							
Correct	70	$\bar{x}$ = 5.83	47	$\bar{x}$ = 3.92	21	$\bar{x}$ = 1.75	138
Incorrect	41	3.42	52	4.33	38	3.17	131
Omitted	9	.75	21	1.75	61	5.08	91
81 - 90							
Correct	71	$\bar{x}$ = 5.92	39	$\bar{x}$ = 3.25	13	$\bar{x}$ = 1.08	123
Incorrect	37	3.08	45	3.75	41	3.42	123
Omitted	12	1	36	3	66	5.5	114
91 - 100							
Correct	60	$\bar{x}$ = 5	39	$\bar{x}$ = 3.25	19	$\bar{x}$ = 1.58	118
Incorrect	42	3.5	44	3.67	34	2.83	120
Omitted	18	1.5	37	3.08	67	5.58	122
TOTAL	1200		1200		1200		3600

4.D.2.3. RESULTS OF CLOZE TEST PASSAGE 4A N = 16

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
1 - 10							
Correct	39	$\bar{x}$ = 7.8	37	$\bar{x}$ = 6.17	29	$\bar{x}$ = 5.8	105
Incorrect	10	2	22	3.67	19	3.8	51
Omitted	1	.2	1	.17	2	.4	4
11 - 20							
Correct	47	$\bar{x}$ = 9.4	44	$\bar{x}$ = 7.33	33	$\bar{x}$ = 6.6	124
Incorrect	3	.6	16	2.67	17	3.4	36
Omitted	-		-		-		-
21 - 30							
Correct	40	$\bar{x}$ = 8	34	$\bar{x}$ = 5.67	26	$\bar{x}$ = 5.2	100
Incorrect	6	1.2	24	4	23	4.6	53
Omitted	4	.8	2	.33	1	.2	7
31 - 40							
Correct	46	$\bar{x}$ = 9.2	39	$\bar{x}$ = 6.5	31	$\bar{x}$ = 6.2	116
Incorrect	3	.6	19	3.17	17	3.4	39
Omitted	1	.2	2	.33	2	.4	5
41 - 50							
Correct	39	$\bar{x}$ = 7.8	29	$\bar{x}$ = 4.83	25	$\bar{x}$ = 5	93
Incorrect	6	1.2	23	3.83	17	3.4	46
Omitted	5	1	8	1.33	8	1.6	21
51 - 60							
Correct	39	$\bar{x}$ = 7.8	40	$\bar{x}$ = 6.67	23	$\bar{x}$ = 4.6	102
Incorrect	8	1.6	18	3	21	4.2	47
Omitted	3	.6	2	.33	6	1.2	11
61 - 70							
Correct	33	$\bar{x}$ = 6.6	37	$\bar{x}$ = 6.17	27	$\bar{x}$ = 5.4	97
Incorrect	11	2.2	15	2.5	18	3.6	44
Omitted	6	1.2	8	1.33	5	1	19

ITEMS	UPPER THIRD	MIDDLE THIRD	LOWER THIRD	TOTAL
71 - 80				
Correct	39 $\bar{x}$ = 7.8	35 $\bar{x}$ = 5.83	28 $\bar{x}$ = 5.6	102
Incorrect	10 2	19 3.17	18 3.6	47
Omitted	1 .2	6 1	4 .8	11
81 - 90				
Correct	39 $\bar{x}$ = 7.8	40 $\bar{x}$ = 6.67	24 $\bar{x}$ = 4.8	103
Incorrect	9 1.8	17 2.83	23 4.6	49
Omitted	2 .4	3 .5	3 .6	8
91 - 100				
Correct	46 $\bar{x}$ = 9.2	45 $\bar{x}$ = 7.5	26 $\bar{x}$ = 5.2	117
Incorrect	3 .6	12 2	15 3	30
Omitted	1 .2	3 .5	9 1.8	13
101 - 110				
Correct	44 $\bar{x}$ = 8.8	36 $\bar{x}$ = 6	31 $\bar{x}$ = 6.2	111
Incorrect	6 1.2	19 3.17	12 2.4	37
Omitted	-	5 .83	7 1.4	12
111 - 115				
Correct	15 $\bar{x}$ = 3	16 $\bar{x}$ = 2.67	6 $\bar{x}$ = 1.2	37
Incorrect	7 1.4	9 1.5	14 2.8	30
Omitted	3 .6	5 .83	5 1	13
TOTAL	575	690	575	1840

4.D.2.4. RESULTS OF CLOZE TEST PASSAGE 5A N = 16

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
1 - 10							
Correct	26	$\bar{x}$ = 5.2	31	$\bar{x}$ = 5.17	17	$\bar{x}$ = 3.4	74
Incorrect	17	3.4	29	4.83	24	4.8	70
Omitted	7	1.4	-		9	1.8	16
11 - 20							
Correct	34	$\bar{x}$ = 6.8	35	$\bar{x}$ = 5.83	20	$\bar{x}$ = 4	89
Incorrect	13	2.6	20	3.33	23	4.6	56
Omitted	3	.6	5	.83	7	1.4	15
21 - 30							
Correct	27	$\bar{x}$ = 5.4	37	$\bar{x}$ = 6.17	25	$\bar{x}$ = 5	89
Incorrect	13	2.6	19	3.17	19	3.8	51
Omitted	10	2	4	.67	6	1.2	20
31 - 40							
Correct	31	$\bar{x}$ = 6.2	35	$\bar{x}$ = 5.83	19	$\bar{x}$ = 3.8	85
Incorrect	15	3	24	4	20	4	59
Omitted	4	.8	1	.17	11	2.2	16
41 - 50							
Correct	31	$\bar{x}$ = 6.2	30	$\bar{x}$ = 5	10	$\bar{x}$ = 2	71
Incorrect	18	3.6	25	4.17	36	7.2	79
Omitted	1	.2	5	.83	4	.8	10
51 - 60							
Correct	39	$\bar{x}$ = 7.8	38	$\bar{x}$ = 6.33	20	$\bar{x}$ = 4	97
Incorrect	9	1.8	21	3.5	26	5.2	56
Omitted	2	.4	1	.17	4	.8	7
61 - 70							
Correct	29	$\bar{x}$ = 5.8	26	$\bar{x}$ = 4.33	8	$\bar{x}$ = 1.6	63
Incorrect	16	3.2	27	4.5	34	6.8	77
Omitted	5	1	7	1.17	8	1.6	20

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
71 - 80							
Correct	31	$\bar{x}$ = 6.2	22	$\bar{x}$ = 3.67	14	$\bar{x}$ = 2.8	67
Incorrect	11	2.2	31	5.17	21	4.2	63
Omitted	8	1.6	7	1.17	15	3	30
81 - 90							
Correct	28	$\bar{x}$ = 5.6	16	$\bar{x}$ = 2.67	9	$\bar{x}$ = 1.8	53
Incorrect	20	4	39	6.5	16	3.2	75
Omitted	2	.4	5	.83	25	5	32
91 - 100							
Correct	27	$\bar{x}$ = 5.4	14	$\bar{x}$ = 2.33	9	$\bar{x}$ = 1.8	50
Incorrect	18	3.6	43	7.17	25	5	86
Omitted	5	1	3	.5	16	3.2	24
TOTAL	500		600		500		1600

4.D.2.5. RESULTS OF CLOZE TEST PASSAGE 2B N = 49

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
1 - 5							
Correct	33	$\bar{x}$ = 2.06	27	$\bar{x}$ = 1.69	12	$\bar{x}$ = .75	72
Incorrect	39	2.44	51	3	57	3.56	147
Omitted	8	.5	7	.41	11	.69	26
6 - 10							
Correct	51	$\bar{x}$ = 3.18	37	$\bar{x}$ = 2.18	27	$\bar{x}$ = 1.69	115
Incorrect	20	1.35	37	2.18	36	2.25	93
Omitted	9	.56	11	.65	17	1.06	37
11 - 15							
Correct	17	$\bar{x}$ = 1.06	4	$\bar{x}$ = .24	2	$\bar{x}$ = .13	23
Incorrect	56	3.5	60	3.53	52	3.25	168
Omitted	7	.44	21	1.24	26	1.63	54
16 - 20							
Correct	28	$\bar{x}$ = 1.75	24	$\bar{x}$ = 1.41	7	$\bar{x}$ = .44	59
Incorrect	39	2.44	46	2.71	43	2.69	128
Omitted	13	.81	15	.88	30	1.88	58
21 - 25							
Correct	30	$\bar{x}$ = 1.88	21	$\bar{x}$ = 1.24	10	$\bar{x}$ = .63	61
Incorrect	45	2.81	55	3.24	52	3.25	152
Omitted	5	.31	9	.53	18	1.13	32
TOTAL	400		425		400		1225

RESULTS OF CLOZE TEST PASSAGE 38 N = 36

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
1 - 5							
Correct	39	$\bar{x}$ = 3.25	37	$\bar{x}$ = 3.08	13	$\bar{x}$ = 1.08	89
Incorrect	17	1.42	16	1.33	35	2.92	68
Omitted	4	.33	7	.58	12	1	23
6 - 10							
Correct	25	$\bar{x}$ = 2.08	20	$\bar{x}$ = 1.67	15	$\bar{x}$ = 1.25	60
Incorrect	27	2.25	33	2.75	32	2.67	92
Omitted	8	.67	7	.58	13	1.08	28
11 - 15							
Correct	43	$\bar{x}$ = 3.58	34	$\bar{x}$ = 2.83	20	$\bar{x}$ = 1.67	97
Incorrect	13	1.08	18	1.5	15	1.25	46
Omitted	4	.33	8	.67	25	2.08	37
16 - 20							
Correct	42	$\bar{x}$ = 3.5	15	$\bar{x}$ = 1.25	4	$\bar{x}$ = .33	61
Incorrect	13	1.08	34	2.83	27	2.25	74
Omitted	5	.42	11	.92	29	2.42	45
21 - 25							
Correct	33	$\bar{x}$ = 2.75	15	$\bar{x}$ = 1.25	5	$\bar{x}$ = .42	53
Incorrect	26	2.17	38	3.17	35	2.92	99
Omitted	1	.08	7	.58	20	1.67	28
TOTAL	300		300		300		900

4.D.2.6. RESULTS OF CLOZE TEST PASSAGE 4B N = 16

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
1 - 5							
Correct	11	$\bar{x}$ = 2.2	13	$\bar{x}$ = 2.17	8	$\bar{x}$ = 1.6	32
Incorrect	11	2.2	15	2.5	17	3.4	43
Omitted	3	.6	2	.33	-		5
6 - 10							
Correct	22	$\bar{x}$ = 4.4	26	$\bar{x}$ = 4.33	9	$\bar{x}$ = 1.8	57
Incorrect	2	.4	4	.67	15	3	21
Omitted	1	.2	-		1	.2	2
11 - 15							
Correct	23	$\bar{x}$ = 4.6	22	$\bar{x}$ = 3.67	18	$\bar{x}$ = 3.6	63
Incorrect	2	.4	7	1.17	7	1.4	16
Omitted	-		1	.17	-		1
16 - 20							
Correct	23	$\bar{x}$ = 4.6	22	$\bar{x}$ = 3.67	10	$\bar{x}$ = 2	55
Incorrect	1	.2	7	1.17	15	3	23
Omitted	1	.2	1	.17	-		2
21 - 25							
Correct	21	$\bar{x}$ = 4.2	16	$\bar{x}$ = 2.67	6	$\bar{x}$ = 1.2	43
Incorrect	4	.8	13	2.17	17	3.4	34
Omitted	-		1	.17	2	.4	3
TOTAL	125		150		125		400

RESULTS OF CLOZE TEST PASSAGE 5B N = 16

ITEMS	UPPER THIRD		MIDDLE THIRD		LOWER THIRD		TOTAL
1 - 5							
Correct	10	$\bar{x}$ = 2	10	$\bar{x}$ = 1.67	5	$\bar{x}$ = 1	25
Incorrect	10	2	16	2.67	18	3.6	44
Omitted	5	1	4	.67	2	.4	11
6 - 10							
Correct	12	$\bar{x}$ = 2.4	11	$\bar{x}$ = 1.83	6	$\bar{x}$ = 1.2	29
Incorrect	13	2.6	18	3	10	2	41
Omitted	-		1	.17	9	1.8	10
11 - 15							
Correct	21	$\bar{x}$ = 4.2	16	$\bar{x}$ = 2.67	6	$\bar{x}$ = 1.2	43
Incorrect	4	.8	14	2.33	5	1	23
Omitted	-		-		14	2.8	14
16 - 20							
Correct	12	$\bar{x}$ = 2.4	11	$\bar{x}$ = 1.83	2	$\bar{x}$ = .4	25
Incorrect	12	2.4	16	2.67	6	1.2	34
Omitted	1	.2	3	.5	17	3.4	21
21 - 25							
Correct	14	$\bar{x}$ = 2.8	11	$\bar{x}$ = 1.83	1	$\bar{x}$ = .2	26
Incorrect	10	2	9	1.5	1	.2	20
Omitted	1	.2	10	1.67	23	4.6	34
TOTAL	125		150		125		400

4.D.3. ITEM ANALYSIS

4.D.3.1. RESULTS OF PASSAGE 2A N = 49

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	necessary	6	3	4	13	.27	.13	
2	the*	16	12	12	40	.82	.25	✓✓
3	would*	11	12	8	31	.63	.19	
4	or*	16	14	7	37	.76	.56	✓✓
5	majority	2	3	0	5	.10	.13	
6	amount	6	3	0	9	.18	.38	
7	the*	16	8	4	28	.57	.75	✓✓
8	as*	6	5	0	11	.22	.38	
9	building	10	4	4	18	.37	.38	✓
10	removed	3	2	0	5	.10	.19	
11	hope	7	8	3	18	.37	.25	✓
12	this*	16	15	12	43	.88	.25	
13	with*	3	2	0	5	.10	.19	
14	cycling	5	6	4	15	.31	.06	
15	you*	13	12	6	31	.63	.44	✓✓
16	liberty	6	3	1	10	.20	.31	
17	later*	15	11	6	32	.65	.56	✓✓
18	without*	15	12	12	39	.80	.19	
19	because*	3	3	2	8	.16	.06	
20	the*	16	15	12	43	.88	.25	
21	irresponsibility	0	0	0	-	-	-	
22	because*	14	10	2	26	.53	.75	✓✓
23	as*	14	14	11	39	.80	.19	
24	individual	9	7	6	22	.45	.19	
25	other	6	6	3	15	.31	.19	
26	rule	13	15	6	34	.69	.44	✓
27	however*	6	3	0	9	.18	.38	
28	is	15	17	14	46	.94	.06	
29	offence	4	5	1	10	.20	.19	
30	any	14	8	6	28	.57	.50	✓
31	made	5	4	0	9	.18	.31	
32	as*	14	14	7	35	.71	.44	✓✓
33	thirteenth	0	0	0	-	-	-	
34	Parliament	14	11	11	36	.73	.19	

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
35	one*	10	7	3	20	.41	.44	✓✓
36	the*	14	9	11	34	.69	.19	
37	the*	8	4	2	14	.29	.38	✓✓
38	read	11	10	6	27	.55	.31	✓
39	and*	14	16	11	41	.84	.19	
40	receive	6	3	2	11	.22	.25	
41	law	15	12	6	33	.67	.56	✓
42	may*	14	15	7	36	.73	.44	✓✓
43	the*	16	15	12	43	.88	.25	
44	the*	12	13	9	34	.69	.19	
45	by*	16	16	13	45	.92	.19	
46	represent	6	4	2	12	.24	.25	
47	applicable	4	2	0	6	.12	.25	
48	the*	11	12	8	31	.63	.25	✓✓
49	the*	14	13	11	38	.78	.25	✓✓
50	courts	6	11	3	20	.41	.25	✓
51	the*	15	11	11	37	.76	.25	✓✓
52	other	3	3	1	7	.14	.13	
53	must*	8	9	5	22	.45	.19	
54	before*	6	5	0	11	.22	.38	
55	given	4	4	0	8	.16	.25	
56	invariably	3	2	0	5	.10	.19	
57	the*	16	17	11	44	.90	.31	
58	and*	7	5	0	12	.24	.44	
59	is	13	12	10	35	.71	.19	
60	and*	8	3	5	16	.33	.19	
61	established	2	0	0	2	.04	.13	
62	trial	6	6	2	14	.29	.25	✓
63	reference	4	0	0	4	.08	.25	
64	sake	4	0	1	5	.10	.19	
65	judge	14	10	6	30	.61	.50	✓
66	called	10	4	4	18	.37	.38	✓
67	was*	9	2	2	13	.27	.44	✓✓
68	involved	2	0	0	2	.04	.13	
69	deliver	6	3	2	11	.22	.25	

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
70	precedent	0	0	0	-	-	-	
71	can*	10	3	2	15	.31	.50	//
72	expression	3	1	0	4	.08	.19	
73	distinguish	1	0	0	1	.02	.06	
74	from*	2	3	4	9	.18	-.13	
75	statute	5	5	0	10	.20	.31	
76	says	10	4	2	16	.33	.50	/
77	intends	2	0	1	3	.06	.06	
78	to*	16	14	12	42	.86	.25	
79	everyone*	4	3	0	7	.14	.25	
80	for*	16	9	6	31	.63	.63	//
81	a*	12	8	3	23	.47	.56	//
82	are*	15	7	8	30	.61	.44	//
83	is	15	13	9	37	.76	.38	/
84	with*	6	1	0	7	.14	.38	
85	this*	7	3	1	11	.22	.38	
86	you*	3	1	0	4	.08	.19	
87	going	6	2	1	9	.18	.31	
88	hearing	7	4	0	11	.22	.44	
89	as*	3	4	0	7	.14	.19	
90	the*	11	10	8	29	.59	.19	
91	if*	15	10	8	33	.67	.44	//
92	why*	6	5	4	15	.31	.13	
93	no	7	7	5	19	.39	.13	
94	and*	15	10	7	32	.65	.50	//
95	law	12	6	3	21	.43	.56	/
96	or*	8	5	5	18	.37	.19	
97	are	15	10	10	35	.71	.31	/
98	shortcomings	2	0	0	2	.04	.13	
99	just	0	0	0	-	-	-	
100	is	7	3	2	12	.24	.31	

ITEM ANALYSIS

4.D.3.1. RESULTS OF PASSAGE 2B N = 49

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	beyond*	0	0	0	-	-	-	
2	tried	2	4	0	6	.12	.13	
3	to*	9	3	4	16	.23	.31	
4	is	13	12	5	30	.61	.50	✓
5	to*	9	8	3	20	.41	.38	✓✓
6	to*	15	9	6	30	.61	.56	✓✓
7	always*	0	1	1	2	.04	-.06	
8	be	14	5	5	24	.49	.56	✓
9	the*	14	16	12	42	.86	.13	
10	with*	8	6	3	17	.35	.31	✓✓
11	not*	6	2	1	9	.18	.31	
12	question	5	1	1	7	.14	.25	
13	facts	4	0	0	4	.08	.25	
14	neighbour's	1	0	0	1	.02	.06	
15	each*	1	1	0	2	.04	.06	
16	prefers	1	1	0	2	.04	.06	
17	disgruntled	1	0	0	1	.02	.06	
18	is	16	13	4	33	.67	.75	✓
19	but*	4	2	1	7	.14	.19	
20	and*	6	8	2	16	.23	.25	
21	actions	0	0	3	3	.06	-.19	
22	to*	15	15	5	35	.71	.63	✓✓
23	situation	5	2	0	7	.14	.31	
24	opinions	8	3	1	12	.24	.44	
25	which*	2	1	1	4	.08	.06	

ITEM ANALYSIS

4.D.3.2. RESULTS OF PASSAGE 3A N = 36

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	possessed	7	4	3	14	.39	.33	✓
2	urban	1	1	0	2	.05	.08	
3	populated	9	7	0	16	.44	.75	✓
4	differentiated	8	5	2	15	.42	.50	✓
5	the*	10	8	2	20	.56	.67	✓✓
6	complex	9	5	3	17	.47	.50	✓
7	labour	9	5	1	15	.42	.67	✓
8	by*	7	4	2	13	.36	.67	✓✓
9	villages	2	2	0	4	.11	.17	
10	development	4	3	0	7	.19	.33	
11	urban	1	0	2	3	.08	-.08	
12	every	0	0	0	-	-	-	
13	in*	10	10	10	30	.83	-	
14	cities	3	1	0	4	.11	.25	
15	and*	11	12	9	32	.89	.17	
16	growth	8	6	2	16	.44	.50	✓
17	is	10	9	7	26	.72	.25	✓
18	the*	10	12	10	32	.89	-	
19	Renaissance	1	3	1	5	.14	-	
20	1500	6	4	4	14	.39	.17	
21	is	11	7	9	27	.75	.17	
22	to*	8	6	4	18	.50	.33	✓✓
23	small	8	3	2	13	.36	.50	✓
24	focal	3	3	0	6	.17	.25	
25	as*	12	9	8	29	.81	.33	✓✓
26	districts	6	4	0	10	.28	.50	✓
27	be	8	7	6	21	.58	.17	
28	two	7	4	1	12	.33	.50	✓
29	geography	3	1	0	4	.11	.25	
30	as*	6	3	0	9	.25	.50	✓✓
31	Europe	8	4	2	14	.39	.50	✓
32	of*	10	7	3	20	.56	.58	✓✓
33	more	11	10	9	30	.83	.17	

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
34	modern	6	2	4	12	.33	.17	
35	size	8	4	0	12	.33	.67	✓
36	s*	10	10	9	29	.81	.08	
37	such*	9	2	0	11	.31	.75	✓✓
38	variety	1	1	0	2	.06	.08	
39	number	3	2	0	5	.14	.25	
40	form	3	1	0	4	.11	.25	
41	accessible	4	3	2	9	.25	.17	
42	pedestrian	5	1	0	6	.17	.42	
43	be*	11	10	9	30	.83	.17	
44	all*	8	6	2	16	.44	.50	✓✓
45	carry	3	1	0	4	.11	.25	
46	were*	11	8	4	23	.64	.58	✓✓
47	stage	10	7	4	21	.58	.50	✓
48	citizens	7	6	1	14	.39	.50	✓
49	towns	8	5	1	14	.39	.58	✓
50	surrounded	9	7	2	18	.50	.58	✓
51	effect	5	4	1	10	.28	.33	✓
52	settlement	8	6	5	19	.53	.25	✓
53	politico-cultural	3	2	0	5	.14	.25	
54	church	0	0	0	-	-	-	
55	a*	6	4	1	11	.31	.42	✓✓
56	sides	8	3	1	12	.33	.58	✓
57	the*	5	2	0	7	.19	.42	
58	on*	12	6	2	20	.56	.83	✓✓
59	easy	5	6	0	11	.31	.42	✓
60	trade	9	4	1	14	.39	.67	✓
61	was	5	1	0	6	.17	.42	
62	two	7	4	0	11	.31	.58	✓
63	a*	9	6	3	18	.50	.50	✓✓
64	ground	2	0	0	2	.06	.17	
65	on*	10	8	3	21	.58	.58	✓✓
66	in*	12	10	7	29	.81	.42	✓✓
67	the*	4	7	3	14	.39	.08	
68	the*	8	7	2	17	.47	.50	✓✓

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
69	or*	2	2	0	4	.11	.17	
70	and*	7	3	3	13	.36	.33	✓✓
71	the*	11	8	5	24	.67	.50	✓✓
72	Ages	5	3	2	10	.28	.25	✓
73	again*	8	7	3	18	.50	.42	✓✓
74	we*	10	8	4	22	.61	.50	✓✓
75	planning	2	1	0	3	.08	.17	
76	basic	5	2	0	7	.19	.42	
77	size	5	2	0	7	.19	.42	
78	know	11	8	4	23	.64	.58	✓
79	up*	5	3	2	10	.28	.25	✓✓
80	only*	8	5	1	14	.39	.58	✓✓
81	of*	12	10	7	29	.81	.42	✓✓
82	to*	8	7	1	16	.44	.58	✓✓
83	firstly	6	0	0	6	.17	.50	
84	factories	1	0	0	1	.03	.08	
85	labour	3	0	1	4	.11	.17	
86	provided	1	1	0	2	.06	.08	
87	of*	12	9	1	22	.61	.92	✓✓
88	in*	10	7	2	19	.53	.67	✓✓
89	of*	10	3	1	14	.39	.75	✓✓
90	in*	8	2	0	10	.28	.67	✓✓
91	provide	5	1	1	7	.19	.33	
92	population	6	2	2	10	.28	.33	✓
93	to*	11	9	2	22	.61	.75	✓✓
94	commerce	2	0	0	2	.06	.17	
95	it*	10	10	6	26	.72	.33	✓✓
96	percent	7	4	2	13	.36	.42	✓
97	towns	10	7	3	20	.56	.58	✓
98	town-dwellers	2	0	0	2	.06	.17	
99	million	4	3	2	9	.25	.17	
100	cities	3	3	1	7	.19	.17	

ITEM ANALYSIS

4.D.3.2. RESULTS OF PASSAGE 3B N = 36

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	complex	8	9	3	20	.56	.42	✓
2	of*	10	12	6	28	.78	.33	✓✓
3	open	10	4	1	15	.42	.75	✓
4	diversified	7	9	3	19	.53	.33	✓
5	administration	4	3	0	7	.19	.33	
6	competing	1	1	0	2	.06	.08	
7	but*	7	10	7	24	.67	-	
8	to*	8	4	3	15	.42	.42	✓✓
9	housing	1	0	0	1	.03	.08	
10	the*	8	5	5	18	.50	.25	✓✓
11	form	2	0	0	2	.06	.17	
12	people	10	9	3	22	.61	.58	✓
13	the*	9	8	5	22	.61	.33	✓✓
14	is	11	5	3	19	.53	.67	✓
15	been*	11	12	9	32	.89	.17	
16	the*	11	5	3	19	.53	.67	✓✓
17	longer	8	2	0	10	.28	.67	✓
18	an*	7	2	0	9	.25	.58	✓✓
19	form	5	1	1	7	.19	.33	
20	own	11	5	0	16	.44	.92	✓
21	that*	12	3	2	17	.47	.83	✓✓
22	by*	9	5	0	14	.39	.75	✓✓
23	future	0	0	1	1	.03	-.08	
24	a*	10	5	1	16	.44	.75	✓✓
25	various	2	2	1	5	.14	.08	

ITEM ANALYSIS

4.D.3.3. RESULTS OF PASSAGE 4A N = 16

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	world	5	5	5	15	.94	-	
2	by*	3	2	1	6	.38	.4	✓✓
3	nesting	4	4	3	11	.69	.2	
4	is	3	6	4	13	.81	-.2	
5	1300	1	0	0	1	.06	.2	
6	these*	5	6	4	15	.94	.2	
7	task	5	1	1	7	.44	.8	✓
8	to*	5	6	5	16	1.00	-	
9	tripod	3	2	1	6	.38	.4	✓
10	the*	5	5	5	15	.94	-	
11	off*	4	6	3	13	.81	.2	
12	the*	5	6	5	16	1.00	-	
13	and*	5	6	5	16	1.00	-	
14	weeks	5	5	4	14	.88	.2	
15	that*	5	4	3	12	.75	.4	✓✓
16	major	4	1	0	5	.31	.8	✓
17	led	4	0	1	5	.31	.6	✓
18	for*	5	4	2	11	.69	.6	✓✓
19	is*	5	6	5	16	1.00	-	
20	for*	5	6	5	16	1.00	-	
21	immediately	3	2	2	7	.44	.2	
22	carapace	5	6	4	15	.94	.2	
23	leatherback	5	4	2	11	.69	.6	✓
24	the*	4	5	3	12	.75	.2	
25	embedded	2	1	2	5	.31	-	
26	bones	2	1	1	4	.25	.2	
27	layer	5	5	5	15	.94	-	
28	thick	5	1	1	7	.44	.8	✓
29	coloration	4	4	3	11	.69	.2	
30	a*	5	5	3	13	.81	.4	✓✓
31	many	5	5	4	14	.88	.2	
32	suggests	5	2	2	9	.56	.6	✓
33	tolerate	4	3	2	9	.56	.4	✓

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
34	its*	5	5	4	14	.88	.2	
35	that*	3	0	0	3	.19	.6	
36	a*	5	6	5	16	1.00	-	
37	the*	5	5	5	15	.94	-	
38	on*	5	6	2	13	.81	.6	✓✓
39	immature	4	1	2	7	.44	.4	✓
40	of*	5	6	5	16	1.00	-	
41	growth	1	0	0	1	.06	.2	
42	bones	4	2	3	9	.56	.2	
43	additional	4	1	0	5	.31	.8	✓
44	that*	5	5	4	14	.88	.2	
45	of*	5	5	4	14	.88	.2	
46	in*	4	4	4	12	.75	-	
47	alone	5	1	1	7	.44	.8	✓
48	in*	5	4	5	14	.88	-	
49	for*	5	5	3	13	.81	.4	✓✓
50	Miami	1	2	1	4	.25	-	
51	last	2	3	2	7	.44	-	
52	of*	5	6	5	16	1.00	-	
53	die	4	5	4	13	.81	-	
54	fit	3	2	2	7	.44	.2	
55	with*	3	3	4	10	.63	-.2	
56	specimens	5	5	1	11	.69	.8	✓
57	have*	4	5	0	9	.56	.8	✓✓
58	their*	5	2	1	8	.50	.8	✓✓
59	our*	4	3	2	9	.56	.4	✓✓
60	and*	4	6	2	12	.75	.4	✓✓
61	the*	5	6	3	14	.88	.4	
62	are*	5	6	5	16	1.00	-	
63	pass	3	3	4	10	.63	-.2	
64	better	3	4	4	11	.69	-.2	
65	was	5	6	4	15	.94	.2	
66	to*	5	6	5	16	1.00	-	
67	back	3	3	1	7	.44	.4	✓
68	limited	0	1	0	1	.06	-	
69	when*	3	1	1	5	.31	.4	✓✓

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
70	nesting	1	1	0	2	.13	.2	
71	into*	2	1	3	6	.38	-.2	
72	it*	5	5	3	13	.81	.4	✓✓
73	exhaustion	5	2	2	9	.56	.6	✓
74	of*	5	6	5	16	1.00	-	
75	another	5	4	4	13	.81	.2	
76	felt	1	1	1	3	.19	-	
77	reach	5	4	2	11	.69	.6	✓
78	flopping	1	0	0	1	.06	.2	
79	of*	5	6	4	15	.94	.2	
80	sea	5	6	4	15	.94	.2	
81	them*	5	6	5	16	1.00	-	
82	source	4	3	1	8	.50	.6	✓
83	because*	5	6	2	13	.81	.6	✓✓
84	in*	3	3	0	6	.38	.6	✓✓
85	their*	5	6	5	16	1.00	-	
86	is*	5	5	3	13	.81	.4	✓✓
87	thus*	3	2	2	7	.44	.2	
88	may*	4	4	1	9	.56	.6	✓✓
89	when*	5	5	5	15	.94	-	
90	anywhere*	0	0	0	-	-	-	
91	of*	5	4	3	12	.75	.4	✓✓
92	tagged	5	4	2	11	.69	.6	✓
93	eggs	4	2	2	8	.50	.4	✓
94	the*	5	6	4	15	.94	.2	
95	the*	5	6	3	14	.88	.4	
96	flow	3	4	2	9	.56	.2	
97	leatherback	5	6	2	13	.81	.6	✓
98	any	4	2	1	7	.44	.6	✓
99	from*	5	6	5	16	1.00	-	
100	for*	5	5	2	12	.75	.6	✓✓
101	far	4	2	2	8	.50	.4	✓
102	in*	5	6	5	16	1.00	-	
103	Guiana	4	4	3	11	.69	.2	
104	Africa	5	2	2	9	.56	.6	✓
105	techniques	3	1	3	7	.44	-	

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
106	of*	5	6	4	15	.94	.2	
107	creatures	5	2	2	9	.56	.6	✓
108	ocean	5	5	3	13	.81	.4	✓
109	beach	3	3	3	9	.56	-	
110	remains	5	5	4	14	.88	.2	
111	in*	3	3	2	8	.50	.2	
112	lead	5	3	0	8	.50	1.00	✓
113	open	5	4	3	12	.75	.4	✓
114	leatherback	1	3	1	5	.31	-	
115	human	1	3	0	4	.25	.2	

ITEM ANALYSIS

4.D.3.3. RESULTS OF PASSAGE 4B N = 16

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	ars	4	6	4	14	.88	-	
2	potentially	2	2	3	7	.44	-.2	
3	what*	0	0	0	-	-	-	
4	called	4	3	0	7	.44	.8	✓
5	all	1	2	1	4	.25	-	
6	leatherback's	5	5	4	14	.88	.2	
7	nests	5	6	0	11	.69	1.00	✓
8	making	3	4	1	8	.50	.4	✓
9	two	4	5	1	10	.63	.6	✓
10	her*	5	6	3	14	.88	.4	
11	cavity	5	2	1	8	.50	.8	✓
12	will*	4	5	5	14	.88	-.2	
13	French	5	4	4	13	.81	.2	
14	the*	5	6	5	16	1.00	-	
15	by*	4	5	3	12	.75	.2	
16	yet*	5	5	2	12	.75	.6	✓✓
17	she*	5	5	2	12	.75	.6	✓✓
18	by*	5	6	3	14	.88	.4	
19	two	3	0	0	3	.19	.6	
20	water	5	6	3	14	.88	.4	
21	renewed	5	3	1	9	.56	.8	✓
22	adapted	4	4	1	9	.56	.6	✓
23	shore	5	5	0	10	.63	1.00	✓
24	recalls	4	3	2	9	.56	.4	✓
25	larger	3	1	2	6	.38	.2	

ITEM ANALYSIS

4.D.3.4. RESULTS OF PASSAGE 5A N = 16

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	portion	0	0	0	-	-	-	
2	information	3	5	0	8	.5	.6	✓
3	actual	1	1	0	2	.13	.2	
4	he*	4	5	5	14	.88	-.2	
5	those*	1	1	0	2	.13	.2	
6	he*	5	6	4	15	.94	.2	
7	the*	4	6	2	12	.75	.4	✓✓
8	social	4	2	3	9	.56	.2	
9	finished	0	0	0	-	-	-	
10	and*	4	5	3	12	.75	.2	
11	predict	5	6	5	16	1.00	-	
12	he*	4	5	2	11	.69	.4	✓✓
13	affect	3	2	1	6	.38	.4	✓
14	artistic	3	3	1	7	.44	.4	✓
15	of*	5	5	3	13	.81	.4	✓✓
16	in*	3	5	4	12	.75	-.2	
17	inventors	5	3	1	9	.56	.8	✓
18	screen	1	0	1	2	.13	-	
19	his*	2	1	0	3	.19	.4	
20	the*	3	5	2	10	.63	.2	
21	raw	2	3	1	6	.38	.2	
22	itself*	0	1	0	1	.06	-	
23	women	1	4	3	8	.5	-.4	
24	with*	5	5	4	14	.89	.2	
25	those*	0	0	0	-	-	-	
26	the*	5	5	5	15	.94	-	
27	experimenting	2	4	3	9	.56	-.2	
28	does*	4	4	2	10	.63	.4	✓✓
29	he*	5	6	3	14	.89	.4	
30	the*	3	5	4	12	.75	-.2	
31	not*	0	2	1	3	.19	-.2	
32	scientific	4	4	1	9	.56	.6	✓
33	their*	4	4	2	10	.63	.4	✓✓

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
34	from*	1	3	2	6	.38	-.2	
35	for*	2	2	2	6	.38	-	
36	whereas*	2	1	2	5	.31	-	
37	history	5	4	2	11	.69	.6	✓
38	in*	5	6	5	16	1.00	-	
39	is	5	3	1	9	.56	.8	✓
40	past	3	6	1	10	.63	.4	✓
41	seeing	4	4	0	8	.5	.8	✓
42	purposes	0	0	0	-	-	-	
43	of*	4	6	5	15	.94	-.2	
44	heavenly	1	0	0	1	.06	.2	
45	step	2	2	0	4	.25	.4	✓
46	alteration	3	3	1	7	.44	.4	✓
47	and*	2	4	0	6	.38	.4	✓✓
48	differs	5	3	1	9	.56	.8	✓
49	changed	5	4	1	10	.63	.8	✓
50	new	5	4	3	12	.75	.4	✓
51	nature	5	3	0	8	.50	1.00	✓
52	strikes	4	4	0	8	.50	.8	✓
53	in*	5	4	3	12	.75	.4	✓✓
54	the*	4	4	1	9	.56	.6	✓✓
55	to*	5	6	5	16	1.00	-	
56	examine	3	3	3	9	.56	-	
57	in*	2	2	2	6	.38	-	
58	of*	5	6	5	16	1.00	-	
59	relation	2	3	0	5	.31	.4	✓
60	animals	4	3	1	8	.5	.6	✓
61	it*	5	6	0	11	.69	1.00	✓✓
62	producers	1	2	1	4	.25	-	
63	utilise	4	4	0	8	.50	.8	✓
64	predecessors	4	3	0	7	.44	.8	✓
65	better	1	0	0	1	.06	.2	
66	fact	3	3	0	6	.38	.6	✓
67	prevented	2	1	0	3	.19	.4	
68	against*	5	5	4	14	.88	.2	
69	a*	1	0	1	2	.13	-	

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
70	their*	3	2	2	7	.44	.2	
71	factor	4	3	0	7	.44	.8	✓
72	analysed	3	1	1	5	.31	.4	✓
73	monkeys	3	0	0	3	.19	.6	
74	the*	4	5	4	13	.81	-	
75	structure	4	1	1	6	.38	.6	✓
76	that*	5	5	3	13	.81	.4	✓✓
77	very*	2	2	0	4	.25	.4	✓✓
78	argues	0	0	1	1	.06	-.2	
79	of*	4	5	4	13	.81	-	
80	once	2	0	0	2	.13	.4	
81	stage	1	2	0	3	.19	.2	
82	man's	4	3	2	9	.56	.4	✓
83	control	4	0	1	5	.31	.6	✓
84	accordingly*	2	2	0	4	.25	.4	✓✓
85	different	3	0	1	4	.25	.4	✓
86	close	2	1	0	3	.19	.4	
87	of*	5	6	4	15	.94	.2	
88	satisfaction	2	0	0	2	.13	.4	
89	enjoyment	1	0	0	1	.06	.2	
90	hence*	4	2	1	7	.44	.6	✓✓
91	if*	3	0	1	4	.25	.4	✓✓
92	valuable	4	1	2	7	.44	.4	✓✓
93	that*	2	0	1	3	.19	.2	
94	the*	5	5	3	13	.81	.4	✓✓
95	nature	0	0	0	-	-	-	
96	in*	3	3	1	7	.44	.4	✓✓
97	point	2	2	0	4	.25	.4	✓
98	competition	1	0	0	1	.06	.2	
99	it*	4	2	1	7	.44	.6	✓✓
100	is	3	1	0	4	.25	.6	✓

ITEM ANALYSIS

4.D.3.4. RESULTS OF PASSAGE 5B N = 16

ITEM NO.	EXACT WORD	UPPER	MIDDLE	LOWER	TOTAL	F.V.	D.I.	COMMENT
1	stage	1	0	3	4	.25	-.4	
2	suffice	0	0	0	-	-	-	
3	command	0	0	0	-	-	-	
4	reality	5	5	0	10	.63	1.00	✓
5	the*	4	5	2	11	.69	.4	✓✓
6	in*	2	1	3	6	.38	-.2	
7	transform	1	0	0	1	.06	.2	
8	possibility	2	4	1	7	.44	.2	
9	ever*	3	3	1	7	.44	.4	✓✓
10	new	4	3	1	8	.50	.6	✓
11	and*	5	6	3	14	.88	.4	
12	other	4	2	0	6	.38	.8	✓
13	transport	2	1	0	3	.19	.4	
14	regions	5	3	1	9	.56	.8	✓
15	ths*	5	4	2	11	.69	.6	✓✓
16	its*	1	1	1	3	.19	.0	
17	being	2	1	0	3	.19	.4	
18	and*	4	3	0	7	.44	.8	✓✓
19	technical	1	2	0	3	.19	.2	
20	of*	4	4	1	9	.56	.6	✓✓
21	itself*	4	3	0	7	.44	.8	✓✓
22	production	1	1	0	2	.13	.2	
23	human	2	0	0	2	.13	.4	
24	the*	5	4	0	9	.56	1.00	✓✓
25	are	2	3	1	6	.38	.2	

4.D.4.1. EXPERIMENT IV : INDIVIDUAL RESULTS N = 49

PASSAGE 2

NUMBER	CLOZE: 2A		CLOZE: 2B		CLOZE: 1	
	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE
1	51	68	4	7	47	68
2	49	63	7	11	44	55
3	48	62	7	11	46	68
4	44	49	5	5	40	48
5	45	53	8	9	32	48
6	44	57	6	11	42	54
7	46	60	6	9	45	64
8	39	50	6	8	35	44
9	37	51	6	9	46	56
10	35	45	7	10	40	49
11	36	45	7	9	39	49
12	29	35	6	8	23	37
13	34	42	5	5	33	38
14	25	32	6	7	27	36
15	38	46	5	6	26	38
16	37	49	7	9	35	47
17	41	54	8	10	24	32
18	37	49	9	10	34	42
19	31	41	4	4	24	29
20	32	43	8	9	27	31
21	27	32	3	4	16	21
22	38	48	5	6	27	39
23	29	35	6	8	27	31
24	33	42	3	6	27	35
25	38	47	4	5	29	40
26	32	40	5	5	24	30
27	24	32	4	4	27	38
28	33	43	6	7	31	39
29	33	47	5	9	30	38
30	23	25	4	5	33	40
31	25	29	4	4	22	26

NUMBER	CLOZE: 2A		CLOZE: 2B		CLOZE: 1	
	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE
32	19	24	4	4	22	26
33	26	37	8	12	21	29
34	26	31	1	1	20	25
35	25	25	4	4	25	31
36	18	23	3	3	15	20
37	31	37	3	3	24	33
38	22	24	6	6	14	19
39	25	31	7	8	22	22
40	17	23	2	2	17	23
41	31	38	6	9	29	38
42	54	68	8	11	36	48
43	26	32	5	6	26	29
44	16	20	2	3	10	17
45	26	30	2	3	15	20
46	21	24	4	5	23	27
47	26	33	6	6	15	18
48	22	26	5	5	14	17
49	36	44	7	9	37	47

4.D.4.2. EXPERIMENT IV : INDIVIDUAL RESULTS N =36

PASSAGE 3

NUMBER	CLOZE: 3A		CLOZE: 3B		CLOZE: 1	
	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE
1	53	67	12	16	48	61
2	52	64	15	16	53	66
3	33	48	12	15	37	44
4	37	45	14	18	41	53
5	55	72	11	14	55	65
6	42	51	10	13	36	52
7	53	67	13	16	40	49
8	42	61	12	14	47	57
9	35	44	10	11	31	42
10	35	45	5	8	21	27
11	35	43	9	13	25	34
12	41	52	7	11	29	38
13	30	41	4	5	29	35
14	29	36	7	7	19	26
15	36	48	14	17	46	62
16	31	38	12	15	26	38
17	23	29	9	9	21	25
18	33	44	10	11	33	45
19	30	46	8	12	27	39
20	41	53	14	15	39	50
21	19	22	6	6	29	37
22	29	31	9	10	26	32
23	32	38	11	13	16	22
24	17	21	6	7	16	20
25	14	16	4	5	14	21
26	20	21	3	4	14	18
27	14	18	7	9	13	14
28	18	22	4	5	16	19
29	22	26	8	9	11	14
30	18	25	7	8	20	24
31	11	15	2	4	10	14

NUMBER	CLOZE: 3A		CLOZE: 3B		CLOZE: 1	
	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE
32	30	36	11	11	21	30
33	4	9	1	2	6	9
34	16	22	2	2	10	13
35	7	9	1	1	18	19
36	20	26	6	8	14	20

4.D.4.3. EXPERIMENT IV : INDIVIDUAL RESULTS N = 16

PASSAGE 4

NUMBER	CLOZE: 4A		CLOZE: 4B		CLOZE: 1	
	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE
1	37	55	9	15	44	55
2	42	68	3	11	44	58
3	53	71	11	17	46	59
4	54	83	9	18	-	-
5	60	89	13	21	-	-
6	81	104	14	22	54	75
7	62	87	9	20	51	71
8	59	90	12	19	63	77
9	71	95	9	17	52	71
10	47	69	5	12	47	60
11	37	52	3	7	56	65
12	39	63	4	9	53	66
13	46	70	7	12	59	72
14	45	71	9	17	56	66
15	45	69	8	18	51	68
16	46	71	9	15	54	65

4.D.4.4. EXPERIMENT IV : INDIVIDUAL RESULTS N = 16

PASSAGE 5

NUMBER	CLOZE: 5A		CLOZE: 5B		CLOZE: 1	
	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE	EXACT	ACCEPTABLE
1	33	48	7	10	41	62
2	38	59	7	12	47	64
3	35	49	8	13	49	68
4	38	51	10	14	54	75
5	29	49	7	8	39	54
6	44	72	10	15	57	79
7	30	45	5	7	53	70
8	32	46	10	15	34	55
9	18	28	2	4	18	20
10	46	69	7	12	48	61
11	29	51	9	12	39	47
12	36	47	7	10	28	40
13	13	20	3	5	27	36
14	28	40	3	4	39	48
15	26	31	5	5	26	37
16	24	33	1	2	31	39

4.E. SAMPLES OF PERFORMANCE

4.E.1. EXPERIMENT II : SUMMARY WRITING

PASSAGE 1

1. The text was about the stages which the human tribe had passed until the stage of civilisation. (1) It shows the improvement of agriculture (2), simple industry (3), and trade (4). Simply shows how the human tribe achieved everything until this period of civilisation.

Marks : 4 P. Kortoutus

2. In the given article was considered the development of mankind from the earliest civilisation like those of Mesopotamia, Siria and Egypt. (1) Development of techniques. (2) We considered the history of civilisation from the earliest stage, when cattle and sheep were used for transport and then horse, camel or ox. (3) We also followed development of agriculture. (4) It was pointed out that tracting plough made revolution in the efficiency of soil processing. (5) It is concluded that labour specialisation improved all kinds of production. (6)

Marks : 6 A. Smailagic

PASSAGE 3A

3. The basic actual large cities are very ancient. (1) About 5000 years before JC the first cities were borne in Egypt, Summaria (2) and Indo Valley.

Initially these cities were surrounded by little villages and they represented a great impulse for the (3) development of cultural activities and economic concentration in that period. Afterwards this model was introduced in Greek and Rome civilisation (4) as a way of make the politics and economy easier. In the West European countries the cities began to appear in the Middle Age. There are 3 stages or periods of the development of the cities. (5) The first one correspond to the Middle Age. since 11 century to until 1500. (6)

The second one to the Renaissance or Baroq Epoc until 1500 and the last one since 18 century until now. (7) In the middle age the first cities had a distinguish focal cultural point, (8) usually a church and from this point city began. In those one where the economic or agricole activities were very important they had a waterway across the city making an enlace with the surrounding villages. (9) Or letting the carry on of produces from a place to another. (10) In the first cities usually there was a distinguish part (the church or sometimes a castle) located in the higher place, a hill for example (11) the rest of the city was in the lowest one giving origin to some special kind of cities. (high-low cities) (12)

For example the Acropolys and certain cities of Belgium and France.

Afterwards the indistrial development (13) with the birth of machines, agricole revolution etc. (14) give origin to the modern cities since 18 century until now.

Marks : 14 R. Arguello

4.E.2. CODE NUMBER: NORW No. 85

CLOZE TEST. ; / INDIVIDUAL PERFORMANCE RECORD # MULTIPLE-CHOICE TEST

NAME: NAT

NATIONALITY THAILAND

TEST SCORE: CLOZE 64 MULTIPLE-CHOICE 12

PROFICIENCY SCORE: DAVIES 42.3 5

ITEM		
1		E
2	DONKEYS	
3	ELEPHANT	A
4	AT	
5		E
6	TIME	
7		E
8		E
9		E
10		E
11		E
12		E
13	-	O
14		E
15	WITH	
16	RECESSIVE	
17		E
18		E
19		E
20	WAS	A
21	SPECIALIZATION	
22	THERE	
23	SHOE	
24	TECHNOLOGICAL	A
25		E
26	PHASE	
27	THE	
28		E
29	THE	
30	IMPROVEMENT	
31		E
32		E
33		E
34	-	O
35		E
36		E
37	TECHNIQUES	A
38		E
39		E
40	AN	A
41		E
42		E
43	1900'S	
44		E
45	INTRODUCTION	A
46	-	O
47		E
48		E
49		E
50	IN	
51	ALSO	A
52	DEVELOPED	A
53	THE	A
54		E
55	SUCH	
56		E
57		E
58	WERE	A
59	SENT	A
60	HERE	
61		E
62	SOME	A
63		E
64	CIVILISED	A
65	AUTHORITY	
66	THERE	
67		E
68	FULLY	
69	COLLECTED	A
70		E
71	THAT	
72	THE	
73	PROGRESS	A
74		E
75	-	O
76	THE	
77		E
78	FORCE	
79		E
80		E
81		E
82	CALENDAR	A
83		E
84	STARS	A
85	IN	
86		E
87		E
88	IN →	E
89	CIVILISATION	A
90	MOST	
91	TO	
92		E
93	IN	A
94	THE	
95	IS	
96	THE	
97	IT	
98	THE	
99	POSSIBLY	
100	THE →	E

SCORE	EXACT	ACCEPTABLE	INCORRECT	OMITTED
100	45	64	32	4

MULTIPLE-CHOICE

VERSION NOS 2

ITEM		
1	b	R
2	d	R
3	a	R
4	c	R
5	c	R
6	a	R
7	b	W
8	c	W
9	c	W
10	a	R
11	c	R
12	b	W
13	b	W
14	c	R
15	a	W
16	c	W
17	d	R
18	c	R
19	a	R
20	c	W
21		
22		
23		
24		
25		

SCORE	RIGHT	WRONG	OMITTED
20	12	8	-

4.E.2. CODE NUMBER: NORW No. 143

CLOZE TEST : INDIVIDUAL PERFORMANCE RECORD + MULTIPLE-CHOICE TEST

NAME : KHADOUDJA BENMBAREK

NATIONALITY MOROCCO - FRENCH

TEST SCORE : CLOZE 62 MULTIPLE-CHOICE 8

PROFICIENCY SCORE : LANG. T. 83 5

ITEM		
1		E
2	ANIMALS →	E
3	NOT	
4		E
5		E
6	PERIOD	
7		E
8		E
9	AGO	
10		E
11		E
12	THE	
13		E
14		E
15	POSSIBILITIES	
16	PRIMITIVE	A
17		E
18		E
19	BY	
20	WAS	A
21	TECHNIQUES	
22	MADE	
23	STONE	A
24	WORK	
25		E
26	TIME	
27		E
28		E
29	ALSO	
30	ADVANCE	
31	THIS	
32		E
33		E
34		E
35		E
36	BASE	
37	SITUATIONS	
38		E
39		E
40		E
41		E
42		E
43	-	O
44		E
45	INVENTION	
46	DO	
47	THEN	
48		E
49		E
50		E
51	IN	
52	BEGAN	A
53	THE	A
54	ECONOMY	
55		E
56		E
57		E
58	WERE	A
59	CARRIED	A
60		E
61	MANUFACTURED	A
62	THE	
63		E
64	LARGE	
65	ONE	
66	-	O
67	ADMINISTRATION	
68		E
69	ESTABLISHED	
70	STARTED	
71		E
72	THE	
73	USE	
74		E
75		E
76	THIS	
77		E
78	TRADE	
79	A	
80		E
81		E
82	CALENDAR	A
83	ONE	A
84	STARS	A
85	THE	A
86		E
87		E
88	IN	
89	POPULATION	A
90		E
91		E
92	WHICH	
93	IN	A
94	BUT → BY	A
95	LED	A
96	THE	
97	TIME	
98	ON	
99	ONLY →	E
100		E

SCORE	EXACT	ACCEPTABLE	INCORRECT	OMITTED
100	46	62	36	2

MULTIPLE-CHOICE

VERSION No: 2

ITEM		
1	b	R
2	d	R
3	d	W
4	d	W
5	a	W
6	a	R
7	b	W
8	c	W
9	d	R
10	a	R
11	a	W
12	d	R
13	c	W
14	d	W
15	b	R
16	d	W
17	d	R
18	a	W
19	b	W
20	c	W
21		
22		
23		
24		
25		

SCORE	RIGHT	WRONG	OMITTED
20	8	12	-

4.E.2. Additional samples of individual performances taken from test materials used in Experiments III and IV.

EXPERIMENT III : CLOZE TEST PASSAGE 1

ITEMS	1 - 10	11 - 20	21 - 30	31 - 40	41 - 50	51 - 60	61 - 70	71 - 80	81 - 90	91 - 100	TOTAL
No. 46 (U)	3 / 7	3 - / 7	6 1 / 3	3 1 / 6	4 1 / 5	4 - / 6	4 1 / 5	3 - / 7	2 - / 8	4 1 / 5	36 5 / 59
No. 52 (L)	4 - / 6	5 - / 5	9 1 / -	6 - / 4	7 1 / 2	4 3 / 3	1 8 / 1	4 2 / 4	5 1 / 4	3 5 / 2	48 21 / 31
No. 2 (U)	2 - / 8	4 - / 6	4 - / 6	1 - / 9	5 - / 5	2 - / 8	5 - / 5	2 - / 8	- - / 10	3 1 / 6	28 1 / 71
No. 26 (L)	4 - / 6	5 1 / 4	6 2 / 2	3 4 / 3	5 3 / 2	4 2 / 4	3 5 / 2	5 3 / 2	3 5 / 2	4 2 / 4	42 27 / 31
No. 30 (M)	4 - / 6	6 - / 4	6 - / 4	4 - / 6	8 - / 2	4 - / 6	7 - / 3	3 - / 7	4 - / 6	6 - / 4	52 - / 48

NOTES: 1. Subjects have been taken from the British Sample and are identified by the number used in their individual result record in Appendix 4.C.3.

2. Central marks are the correct responses by acceptable-word method. Marks to the top left are incorrect responses, and marks at top right of each box are the number of omitted responses.

3. (U), (M), (L) indicate, upper, middle or lower third category divided on the basis of total scores.

EXPERIMENT IV

No. 1 (U)	4 1 5	3 1 6	2 - 8	4 - 6	3 - 7	2 1 7	2 1 7	1 1 8	3 - 7	2 2 6	26 7 67
No. 9 (M)	5 - 5	4 - 6	6 - 4	8 - 2	4 - 6	5 - 5	6 - 4	8 - 2	3 1 6	6 - 4	55 1 44
No. 30 (L)	3 5 2	8 - 2	3 4 3	5 2 3	3 3 4	4 6 -	6 3 1	3 2 5	3 5 2	6 1 3	44 31 25

PASSAGE 3A (Appendix 4.D.4.2.)

No. 10 (U)	2 4 4	1 1 8	2 2 6	2 1 7	2 1 7	2 - 8	2 - 8	2 1 7	2 - 8	4 - 6	21 10 69
No. 7 (M)	5 - 5	3 2 5	4 1 5	3 - 7	4 3 3	3 - 7	3 2 5	5 1 4	8 1 1	6 1 3	44 11 45
No. 13 (L)	4 2 4	4 1 5	7 - 3	6 3 1	8 1 1	5 - 5	6 3 1	- 10 -	- 10 -	- 10 -	20

PASSAGE 5A (Appendix 4.D.4.4.)